

Proceedings of:

The Eleventh Annual Meeting

The American Society for Precision Engineering

November 9-14, 1996

*Monterey Conference Center and Doubletree Hotel at Fisherman's Wharf
Monterey, California*

The American Society for Precision Engineering (ASPE) is a multidisciplinary professional and technical society concerned with research and development, design, manufacture and measurement of high accuracy components and systems. ASPE activities encompass relevant aspects of mechanical, electronic, optical and production engineering, physics, chemistry, and computer and materials science. Membership is open to anyone interested in any aspect of precision engineering.

Founded in 1986, ASPE provides a new focus for a diverse but important community. Other professional organizations have covered aspects of precision engineering, always as a sideline to their principal goals. ASPE is based on the core of generic concepts necessary to achieve precision in any application; independent of discipline, ASPE intends to be the focus for precision technology — and to represent all facets from research to application.

*ASPE — The American Society for Precision Engineering
401 Oberlin Road, Suite 108
P.O. Box 10826
Raleigh, NC 27605-0826
Phone (919) 839-8444 Fax (919) 839-8039*

Preface

This book comprises the proceedings of the 1996 Annual Meeting. The contributions reflect the authors' opinions and are published as presented to ASPE, without change. Their inclusion in this publication does not necessarily constitute endorsement by the ASPE, or its editorial staff.

Individual readers of this book and nonprofit libraries acting for them are permitted to make fair use of the material contained in this book to be used in teaching or research; provided they copy said paper in total, as written by the author. Permission is granted to quote excerpts from the scientific or technical works in this book with proper acknowledgement of the source to include: the author's name, the book name, and the ASPE tag line. Reproduction of figures and tables is likewise permitted from the contents of this book, provided that the same acknowledgment of the same source and content is printed.

Any other type reproduction, systematic copying or multiple reproduction of any material in this book is expressly prohibited except with the written permission of the ASPE editorial staff, addressed to the Society's headquarters.

ISBN 1-887706-16-X

1996

*American Society for Precision Engineering
401 Oberlin Road, Suite 108
Raleigh, NC 27605-0826*

Printed in the United States of America

Welcome Note

The American Society for Precision Engineering celebrates its Eleventh Annual Meeting this year in Monterey, California. The remarkable growth and success of the Society is reflective of the increasing importance of precision engineering in a wide variety of fields of endeavor, from manufacturing to microelectronics to basic science. The Society serves as a focus for precision engineers across all of these fields. The Annual Meeting has evolved into a premier international forum for the exchange of ideas and presentation of research results relating to precision engineering problems involving mechanical parts, electronic components, materials and material processing, metrology, controls, and system integration. Precision engineers and scientists from private industry, government laboratories, and universities meet to learn about the latest developments and to exchange ideas about the future directions of these technologies.

This year's meeting continues the tradition of offering two full days of tutorials presented by some of the foremost precision engineers and scientists in the world. Many participants cite the tutorial program as one of the most valuable features of the meeting. The technical sessions and poster presentations offer the latest research results in the areas of ultraprecision machining, metrology, controls, precision grinding, micro-positioning, material processing, design, precision transducers, surface profilometry, machine tool metrology, and error compensation. The commercial exhibit will provide attendees with the opportunity to view and discuss the latest precision engineering equipment, products, and services that are commercially available. The open forum will give attendees the opportunity to express their opinions and pose questions for discussion by the group, enhancing the exchange of information.

The conference organizing committee is proud to present the program for the Eleventh Annual Meeting of ASPE. We welcome your participation.

Welcome to Monterey.

1996 ASPE Meeting Organizing Committee

*Robert G. Wilhelm, Chairman
University of North Carolina at Charlotte*

*Jeffrey W. Carr, IBM Corporation
Kirsten M. Carr, Ford Research Laboratory
James F. Cuttino, University of Alabama
Thomas A. Dow, North Carolina State University
Joseph D. Drescher, Pratt & Whitney
Anthony E. Gee, Cranfield University
David Grigg, Digital Instruments
Andrew J. Hazelton, Sandia National Laboratory
Lowell P. Howard, Nat'l Institute of Standards and Technology
Jang Moo Lee, Seoul National University
Don A. Lucca, Oklahoma State University
Dan E. Luttrell, 3M Company
Michele M. Miller, Michigan Technological University
Robert Pilkey, Sandia National Laboratory
Stuart T. Smith, University of North Carolina at Charlotte
Toyama Shigeki, Tokyo A&M University
H. Frank van Beek, Delft University of Technology
Hong-Tzong Yau, National Chung Chen University*

Program

1996 ASPE Annual Meeting

Time	Sat., Nov. 9	Sun., Nov. 10	Mon., Nov. 11	Tues., Nov. 12	Wed., Nov. 13	Thurs., Nov. 14
8:00			Continental Breakfast			
9:00	Tutorials	Tutorials	Technical Session 1	Technical Session 4	Technical Session 7	Technical Tours
10:00			Break	Break	Break	
11:00			Technical Session 2	Poster Session 1 + Exhibits	Poster Session 2 + Exhibits	
12:00					Lunch + Committee Meetings	
1:00						
2:00	Tutorials	Tutorials	Commercial Session	Technical Session 5	Technical Session 8	
3:00			Break	Break	Break	
4:00			Technical Session 3	Technical Session 6	Technical Session 9	
5:00				Hospitality Suite		
6:00			Hospitality Suite	Keynote Address		
7:00		Welcome Reception				
8:00			Dinner	Open Forum		
9:00						

1996 Annual Meeting

Contents

<i>Preface</i>	<i>ii</i>
<i>Welcome Note—Organizing Committee</i>	<i>iii</i>
<i>Program</i>	<i>iv</i>
<i>Sponsor Members</i>	<i>vi</i>
<i>Cooperating Sponsors—Technical Exhibitors</i>	<i>vii</i>
<i>Keynote Address/Forum/Commercial</i>	<i>1</i>
<i>Technical Paper Index</i>	<i>2-13</i>
<i>Technical Papers</i>	<i>15-675</i>
<i>Authors Index</i>	<i>677-680</i>

Sponsor Members of the ASPE

3M Company

3M Center, Bldg. 235-1S-46
St. Paul, MN 55144
(612) 575-1465
Contact: William J. Bryan

ADE Technologies, Inc.

80 Wilson Way
Westwood, MA 02090
(617) 467-3600
Contact: John Morgello

Bausch & Lomb

1400 North Goodman Street
Rochester, NY 14692-0450
(716) 338-6754
Contact: Michael J. Clark

Brown & Sharpe

Precision Park
North Kingstown, RI 02852
(401) 886-2358
Contact: David H. Genest

Chardon Tool & Supply Company, Inc.

P.O. Box 291
11993 Ravenna Road #18
Chardon, OH 44024
(216) 286-6440
Contact: Weldon Bennett

Contour Fine Tooling, Inc.

P.O. Box 111, Fisk Mill
7 Water Court
Marlborough, NH 03455
(603) 876-4026
Contact: Jeffrey Tomlin

Cranfield Precision, Div. of Landis Lund

A Western Atlas Corporation
90 Central Avenue, Wharley End
Cranfield, Bedford MK43 0JR, England
44-1234-750111
Contact: Keith G. Dobson

Dover Instrument Corporation

P.O. Box 200
200 Flanders Road
Westboro, MA 01581
(508) 366-1456
Contact: John R. Hero

Fabreeka International, Inc.

1023 Turnpike Street
Stoughton, MA 02072
(617) 341-3655
Contact: Robert P. Haley

Federal Products Company

1144 Eddy Street
P.O. Box 9400
Providence, RI 02940
(401) 784-3403
Contact: Frederick G. Parsons

Lawrence Livermore National Laboratory

P.O. Box 808, L-644
Livermore, CA 94550
(510) 422-1915
Contact: Daniel C. Thompson

Moore Special Tool Company

800 Union Avenue
Bridgeport, CT 06607-0088
(203) 366-3224
Contact: Wayne F. Hobson

Precitech, Inc.

P.O. Box 409
Keene, NH 03431
(603) 357-2510
Contact: Donald Brehm

Professional Instruments Company

4601 Highway 7
Minneapolis, MN 55416
(612) 927-4494
Contact: Ted J. Arneson, Jr.

Reflexite-Precision Technology Center

200 Park Centre Drive
West Henrietta, NY 14586
(716) 272-0290
Contact: Steven M. Scott

Taylor Hobson / Pneumo

Precision Park, P.O. Box 543
Keene, NH 03431
(603) 352-5242
Contact: Leonard E. Chaloux

Tencor Instruments

3333 Octavius Drive
Santa Clara, CA 94054
(408) 970-9500
Contact: Graham J. Siddall

Tropel Corporation

60 O'Connor Road
Fairport, NY 14450
(716) 388-3469
Contact: Louis Denes

Zygo Corporation

Laurel Brook Road
P.O. Box 448
Middlefield, CT 06455-0448
(203) 347-8506
Contact: Carl A. Zanon

Cooperating Sponsors

Elsevier Science, Inc.

Photonics Spectra

Technical Exhibitors

ADE Technologies, Inc.

Bal-tec Division

Chardon Tool & Supply Co., Inc.

Contour Fine Tooling, Inc.

Cranfield Precision, Division of Landis Lund

Dover Instrument Corporation

Fabreeka International, Inc.

Heidenhain Corporation

Lion Precision

National Precision Manufacturing Institute

Phase Shift Technology

Precision Engineering Center—N.C. State University

Precision Engineering Laboratory—University of N.C. at Charlotte

Professional Instruments Co., Inc.

Queensgate Instruments

Surface Finishes

Taylor Hobson / Pneumo

Tetra Precision, Inc.

The Timken Company

WYKO Corporation

Zygo Corporation

Keynote Address

E. Michael Campbell

Tuesday, November 12, 6:00 p.m.

Dr. E. Michael Campbell, Associate Director for Laser Programs at the Lawrence Livermore National Laboratory, will deliver the Keynote Address on "Precision Engineering Frontiers in Inertial Confinement Fusion." This technology is key to the success of the National Ignition Facility (NIF), a planned U.S. Department of Energy project where energy from an extremely powerful laser will be used to "ignite" small capsules filled with fusion fuel, creating a fusion reaction and liberating more energy than was required to start the process. With applications and spin-offs in power generation, optics, micro-electronics, basic science research, and defense, the NIF will entail a very precise and large-scale engineering effort well into the next century.

As Associate Director for Laser Programs, E. Michael Campbell is responsible for the overall program management and technical development of the multidisciplinary Laser Directorate. The major activities of this organization include the Inertial Confinement Fusion Program, Isotope Separation, Advanced Manufacturing Program, Advanced Microtechnology Program, Imaging and Detection Program, and the Remote Sensing Project, as well as the active development of promising new technologies in the areas of lasers and electro-optics and microfabrication.

Prior to his current position, Campbell was Deputy Associate Director and Program Leader for the Inertial Confinement Fusion Program. Campbell is a member and Fellow of the American Physical Society, Division of Plasma Physics. He has participated in National Academy panels on plasma, laser, and optical science. He is a recipient of the 1985 U.S. Department of Energy's Excellence in Weapons Research Award for the development of the x-ray laser, the 1990 American Physical Society Award for Excellence in Plasma Physics Research, and the Edward Teller Medal in 1994. Most recently, Campbell was honored with the 1994 E. O. Lawrence Award, presented for distinguished leadership in helping to propel laser-driven inertial confinement fusion to the forefront of physics research, and the Fusion Power Associates 1995 Leadership Award. He received his B.S. in Mechanical Engineering from the University of Pennsylvania, followed by an M.A. and Ph.D. in Applied Physics, Aerospace and Mechanical Sciences from Princeton University.

Commercial Session

Monday, November 11, 1:30 p.m.

Session Chair: Robert G. Wilhelm, University of North Carolina at Charlotte

The commercial session will take place on Monday afternoon, in a special session where company representatives are invited to make brief presentations on new products associated with precision engineering. Conference participants will have the opportunity to hear about limitations of performance of existing approaches, new approaches relevant to precision engineering, and advances that can be expected in the future.

Open Forum

Tuesday, November 12, 7:00 p.m.

Session Chair: Robert J. Hocken, University of North Carolina at Charlotte

The open forum was created to encourage dissemination of new and significant results, as well as observations on precision engineering subjects. All interested conference participants can sign up at the Conference Registration Table for short, informal presentations.

Technical and Poster Paper Index

The technical program for the 1996 ASPE Annual Meeting contains over 140 papers on current research from Asia, Australia, Europe, and South and North America. Authors from national laboratories, universities and industry will share their ideas and conclusions on a variety of topics.

Papers that lent themselves to significant verbal interaction, concise but important discoveries, or strongly visual or tactile subjects have been selected for the poster session. This year the poster papers will be close to the exhibits for the duration of the conference, and authors will be present to discuss their work on Tuesday, November 12 and Wednesday, November 13, from 10:30 a.m. to 12:00 p.m.

Once again, in addition to ten technical sessions, an Open Forum and a Commercial Session will be offered.

Session I

Nanometer Positioning

Monday, November 11, 1996

8:30 a.m.- 10:00 a.m.

Session Chair: H. Frank van Beek (Delft University of Technology)

1. **Positioning Performance of a Maglev Fine Positioning System**
J. B. Wronosky, T. G. Smith, J. D. Jordan, J. R. Darnold (Sandia National Laboratories) page 15
2. **Design and Characterization of Nanometer Precision Mechanisms**
Y. Xu, P. D. Atherton, M. McConnell, T. R. Hicks (Queensgate Instruments Ltd.) page 20
3. **A Piezo-driven TEM Stage with Subnanometer Position Stability**
H. van der Wulp, H. F. van Beek, B. M. Mertens (Delft University of Technology) page 26
4. **Ultraprecise Traveling Slider for Fabrication of DVD Glass Master**
J. Nishida, N. Kikuri, N. Uchida, N. Takasu (Toshiba Corporation) page 32

Session II

Nanometer Instrumentation

Monday, November 11, 1996

10:30 a.m. - 12:00 p.m.

Session Co-chairs: Anthony E. Gee (Cranfield University) and Andrew J. Hazelton (Nikon Research Corporation of America)

1. **An Electrodynamic Microactuator for the Aligning of an Optical Fibre with an Integrated Position Sensor**
T. Frank (Technical University of Ilmenau) page 36
2. **Observations of Contact Measurements Using a Resonance Based Touch Sensor**
M. Vidic, S. M. Harb (Coventry University), and S. T. Smith (University of North Carolina at Charlotte) page 41

3. **Novel Linear Encoder Incorporating Incremental Optical and Absolute Capacitive Encoders**
S. A. Hogan, K. G. Masreliez (Micro Encoder Inc.) page 47
4. **A New Design for a High Speed Spindle**
M. Weck, S. Fischer (Fraunhofer-Institut für Produktionstechnologie, IPT), P. Holster
(Philips Research Labs), K. Carlisle (Cranfield Precision Ltd.), and Y. Chen (Carl Zeiss) page 51

Session III

Ultraprecision Machining

Monday, November 11, 1996

3:30 p.m. - 5:00 p.m.

Session Chair: Richard L. Rhorer (National Institute of Standards and Technology)

1. **Precision Grinding of Brittle Materials: The ASPE 1996 Topical Meeting—
Summary and Review (Invited Talk)***
R. O. Scattergood (North Carolina State University) page 55
2. **Diamond Machining of the 3m Reflector of the KOSMA Submillimeter Telescope by a
Single-point Fly-cutting Process**
E. Brinksmeier, W. Preuss (University of Bremen) page 56
3. **The Effect of Spindle Perturbations on the Quality of Precision Grinding**
K. L. Blaedel, J. E. Bibler, J. S. Taylor (Lawrence Livermore National Laboratory) page 62
4. **Possible Mechanism of Crack Initiation in Micromachining of Silicon**
S. Shimada (Osaka University), T. Inamura, N. Takezawa (Nagoya Institute of Technology),
N. Ikawa (Osaka Electro-Communication University), Y. Onchi (Mizuho Co., Ltd.), and T. Sata
(Toyota Technological Institute) page 65

Session IV

Novel Processing

Tuesday, November 12, 1996

8:30 a.m. - 10:00 a.m.

Session Chair: Michele H. Miller (Michigan Technological University)

1. **Precision Manufacturing of Compact Disc Master Stampers**
T. G. Bifano, H. Caggiano (Boston University), and P. Bierden (Prism Corporation) page 69
2. **A Tabletop Factory to Fabricate Micro Machines, "Nano Manufacturing World"**
M. Nakao, Y. Hatamura (The University of Tokyo) page 74
3. **In-situ Slurry Film Measurements for Chemical-mechanical Polishing**
J. Levert, F. Mess, S. Danyluk, R. Salant (Georgia Institute of Technology), and
R. Baker, L. Cook (Rodel, Inc.) page 80
4. **Three-dimensional Profile Generation by Dot-matrix Method Using Miniaturized
Electrode Feeding Devices**
K. Furutani, T. Enami, N. Mohri (Toyota Technological Institute) page 86

*Extended abstract not available at press time.

Poster Session I

Tuesday, November 12, 1996

10:30 a.m. - 12:00 p.m.

NEW METROLOGY DEVICES

1. **Nanometric Measurement by Optical Interferometry**
M. J. Downs (National Physical Laboratory) page 91
2. **Study on a Novel Method for Refractive Index Compensation of Laser Measuring System on Ultraprecision Machine Tools**
F. Z. Fang, Y. Q. Zhang, G. X. Zhang (Tianjin University) page 95
3. **Determination of Straight Line Profiles by Using Angularly Displaced CCD with Pixel Output Interpolation**
R. Jablonski (Shizuoka University and Warsaw University of Technology) page 100
4. **The Development of Non-contact 3D Profile Measurement Technique with Parallel Projection Beams**
S-J. Lee, T-R. Fon (Yuan-Ze Institute of Technology) page 104
5. **A Method to Analyze Zone-plate Interferograms by Using Nematic Liquid Crystal**
T. Nomura, K. Kamiya, H. Miyashiro (Toyama Prefectural University), and K. Yoshikawa, H. Tashiro, H. Onnagawa, H. Okada (Toyama University) page 110
6. **Straightedge Metrology in a Wet Environment**
S. B. Peirce, D. B. DeBra (Stanford University) page 114
7. **Machining Virtual Test Parts**
T. Schmitz, J. C. Ziegert (University of Florida) page 118
8. **A Microfabricated Magnetic Field Sensor Based on Electrodeposited Giant Magnetoresistant Material**
W. Wang, A. Natarajan, E. Ma, C. Khan-Malek (Louisiana State University) page 122
9. **Error Analysis by Simulation for a Fiber Based Non-contact Measurement Probe System**
Y. Yang, K. Yamazaki (University of California at Davis) page 126
10. **Micro-volume Defects Measuring System Based on Digital Image Processing Measuring Technique**
You, Z., Yang, R., Chen, J. (Tsinghua University) page 132
11. **A Study on High-accuracy Laser Interferometric Measurement of Micro-angle**
Zhu, H., Wang, H., Liu, H. (Tianjin University) page 136

SURFACE ROUGHNESS

1. **Stylus Drag Forces on Lubricated Surfaces**
D. G. Chetwynd, L. A. J. Davis (University of Warwick) page 141
2. **Development of Dual Servo STM**
K. Mitsui (Keio University), and T. Goto (Mitsubishi Heavy Industries, Ltd.) page 146

SURFACE SHAPE AND FIGURE

1. **Direct Comparative Length Measurement Between Regular Crystalline Lattice and SEM Standard Grating Using Dual Tunneling Unit STM**
M. Aketagawa, K. Takada, S. Sasaki, S. Suzuki, K. Kobayashi, K. Yamada (Nagaoka University of Technology), and Y. Nakayama (Hitachi Ltd.) page 150
2. **Cylindricity — More Than Just a Number?**
P. J. Castle (Rank Taylor Hobson, Inc.) page 154
3. **Optical Polygon Calibration with a He-Ne Laser**
R. Díaz, R. Nava (Universidad Nacional Autónoma de México) page 158

4. Nano-metrology of Roundness and Spindle Error W. Gao, S. Kiyono, T. Sugawara (Tohoku University)	page 163
5. Spindleless Instrument for the Roundness Measurement of Precision Balls E. A. Gleason, Jr., H. Schwenke (Bal-Tec, Division of Micro Surface Engineering)	page 167
6. Assessment of Drill Point Geometry by Genetic Algorithm and Steepest Gradient Method L. Hazra, R. Nakamura, H. Kato (Chiba University), T. Kuroda (Kisarazu National College of Technology), and Y. Tsuchiya, I. Sakuma (Fujita Works)	page 172
7. Transmission Electron Microscopy of Precision Machined Surfaces T. Kaneeda, D. Higashi (Okayama University of Science), and S. Yokomizo (Industrial Technology Center of Okayama Prefecture)	page 176
8. Statistical Methods for Process Identification by Surface Classification Using Vision System M. B. Kiran, B. Ramamoorthy, V. Radhakrishnan (Indian Institute of Technology)	page 180
9. Highly Accurate Technique for SPM Measurement S. Kiyono, M. Satoh, T. Nanjo, W. Gao (Tohoku University)	page 186
10. A 120mm x 120mm Area Photomask Metrology Standard with Absolute Position Accuracy of 60nm M. B. McCarthy (National Physical Laboratory), and A. E. Gee (Cranfield University)	page 190
11. An Optical Calibrating Technique for Non-contact Metrology M. T. Mostafavi, S. Pragalsingh (University of North Carolina at Charlotte)	page 196
12. Flatness and Levelness Measurement System for Big Surfaces R. Nava, S. Padilla, B. Valera (Universidad Nacional Autónoma de México)	page 202
13. New NIST-Certified Length Microscale J. Potzick (National Institute of Standards and Technology)	page 206
14. Height Calibration of Atomic Force Microscopes Using Silicon Atomic Step Artifacts V. W. Tsai, E. D. Williams (University of Maryland at College Park), T. V. Vorburger, P. J. Sullivan, R. Dixon, R. Silver, J. Schneir (National Institute of Standards and Technology)	page 212
15. Evaluating Form Error of Spherical Surfaces Using Confocal Microscope G. Udupa, M. Singaperumal, R. S. Sirohi, M. P. Kothiyal (Indian Institute of Technology)	page 217

COORDINATE METROLOGY MACHINES

1. Thermal Errors Calibration and Modeling of a Cylindrical Grinding Machine CNC B. DiGiacomo (Universidade de Sao Paulo), P. H. Pereira (University of North Carolina at Charlotte), and A. J. Abackerli (Universidade Metodista de Piracicaba)	page 223
2. Substitute Geometry Fitted from Position and Surface Normal Information R. Edgeworth, R. G. Wilhelm (University of North Carolina at Charlotte)	page 229
3. Modeling and Simulation of a Parallel Architecture, Five-axis Coordinate Measuring Machine B. Jokiel, Jr., J. C. Ziegert (University of Florida)	page 233
4. Evaluation of Effect of Sample Point Density on Dimensional Measurements of Cylindrical Features Using Coordinate Measuring Machine A. C. Okafor, D. Mehrotra (University of Missouri-Rolla)	page 237
5. The Effect of Circular Groove on Flatness Error of Metal Mirror Caused by Chucking Force N. Ozawa, K. Mizuhara (Mechanical Engineering Laboratory, AIST, MITI), and S. Kawabe (Precision Technology Research Institute of Nagano Prefecture)	page 243
6. Sampling Plan for Point-to-point Performance Test of Touch Trigger Probes on Coordinate Measuring Machines Y-L. Shen, S. Moon (George Washington University)	page 247
7. A Robust Pretravel Model for Touch Trigger Probes in Coordinate Metrology Y-L. Shen, M. E. Springer (George Washington University)	page 253
8. Calibration Methods for Dual Probe CMMs B. P. Vishwas, R. G. Wilhelm (University of North Carolina at Charlotte)	page 259

9. **Evaluation of Elliptical Profiles and its Application in Cylinder Measurement by CMMs**
Y. Wang, K. S. Moon, J. W. Sutherland (Michigan Technological University) page 263
10. **Investigation of Error Motions of Aerostatic Spindles**
M. Weck, M. F. Soltani (Fraunhofer-Institut für Produktionstechnologie, IPT) page 267
11. **Surface Approximation and Generation of Complex Geometric Model from Laser Scan Data**
H-T. Yau, J-S. Chen (National Chung Cheng University) page 273

MACHINE TOOL METROLOGY

1. **Machine Tool Accuracy Characterization Using Telescopic Ballbar**
A. C. Okafor, Y. M. Ertekin (University of Missouri-Rolla) page 278
2. **A New Technique for Volumetric Error Assessment of CNC Machine Tools Incorporating the Ballbar Measurement and the Three-dimensional Volumetric Error Model**
H. J. Pahk, Y. S. Kim, J. H. Moon (Seoul National University) page 284
3. **Analysis of Tool Trajectories (2D Tool Paths) in Machine Tools**
F. Rehsteiner, S. Weikert (Swiss Federal Institute of Technology) page 288

DESIGN I

1. **Comparison of Control Strategies for Piezoelectric and Magnetostrictive Micropositioning Devices for Ultraprecision Machining**
J. C. Campos Rubio, J. G. Duduch, A. J. Vieira Porto, H. Lêdo (Universidade de São Paulo), and A. E. Gee (Cranfield University) page 294
2. **A Study on Feeding Control Mechanism Analysis of Industrial Lock Stitch Sewing Machine**
K. J. Chun, D. Y. Shin (Korea Academy of Industrial Technology) page 300
3. **Piezoelectric Microactuator Including Compliant Mechanism**
D. Enke (Technical University of Ilmenau) page 306
4. **A Semi-toroidal Flexure System for Extended Linear Motion**
C. J. Finlayson, J. F. Cuttino (The University of Alabama) page 310
5. **Microscopic Behavior of Leadscrew-nut System and its Improvement**
S. Fukada (Shinshu University) page 314
6. **Recent Advancements in Piezoelectric Stepping Motors**
D. Henderson, D. Jensen, P. Piccirilli (Burleigh Instruments, Inc.) page 318
7. **A Long-range Scanning Stage Design (The LORS Project)**
M. L. Holmes, R. J. Hocken (University of North Carolina at Charlotte), and D. L. Trumper (Massachusetts Institute of Technology) page 322
8. **A Study on a High-precision Linear Motion Table with Magnetic Bearing Suspension**
Y. Joung (Samsung Electronics Co.), H-J. Ahn, O-S. Kim, D-C. Han (Seoul National University), and I-B. Chang (Kangwon National University) page 328
9. **Fabrication and Thermal Analysis of a Gas-liquid Phase-change Micro-actuator**
S. Kato (Nippon Institute of Technology) page 332
10. **Load-dependent Behaviour in PTFE-glass Slideways**
A. F. Monteiro, D. G. Chetwynd (University of Warwick) page 336
11. **Micro-dynamics of a Slide Mechanism Using Rolling Elements**
Y. Okazaki (Mechanical Engineering Laboratory, AIST), and S. Sasaki, R. Murata (Science University of Tokyo) page 342
12. **Characterization of an Air Bearing for Mixing Studies of Bioprocesses**
J. M. Patiño, R. Vázquez, G. Ascanio, E. Galindo, L. Serrano-Carreón (Universidad Nacional Autónoma de México) page 346

Session V

Advances in Interferometry

Tuesday, November 12, 1996

1:30 p.m. - 3:00 p.m.

Session Co-chair: Michael J. Downs (National Physical Laboratory) and Lowell P. Howard (National Institute of Standards and Technology)

1. **Industrial Inspection Interferometer Having a Large Equivalent Wavelength (Invited Paper)**
P. de Groot, C. Zanoni (Zygo Corporation) page 350
2. **Application of Liquid Crystal Panel as Diffraction Grating to Zone-plate Interferometer**
K. Kamiya, T. Nomura, H. Miyashiro (Toyama Prefectural University), and K. Yoshikawa, H. Tashiro (Toyama University) page 353
3. **Wavelength-shift Interferometry: Using a Dither to Improve Accuracy**
J. A. Stone, L. P. Howard (National Institute of Standards and Technology), and A. Stejskal (Institute of Scientific Instruments, AVCR) page 357
4. **Angle Measurement Using X-ray Moiré Fringes**
N. O. Krylova, D. G. Chetwynd, D. K. Bowen (University of Warwick) page 363

Session VI

Nanometrology

Tuesday, November 12, 1996

3:30 p.m. - 5:00 p.m.

Session Chair: David Grigg (Digital Instruments Inc.)

1. **Scanning Probe Microscopes for Dimensional Length Measurements**
K. Hasche, O. Jusko, L. Koenders, K. Thiele, G. Wilkening, X. Zhao (Physikalisch-Technische Bundesanstalt) page 369
2. **Toward Accurate Measurements of Pitch, Height, and Width Artifacts with the NIST Calibrated Atomic Force Microscope**
R. Dixon, T. V. Vorburger, P. J. Sullivan, T. H. McWaid (National Institute of Standards and Technology), and V. W. Tsai (University of Maryland at College Park) page 375
3. **Measurements of the Surface Roughness of Ground Silicon Wafers by Cross-sectional TEM**
H. Murai, S. Kasaishi, T. Kaneeda (Okayama University of Science), Y. Yokota (Natural Science Research Institute), and P. O. Hahn, M. Kerstan, D. Helmreich (Wacker Silitronic) page 381
4. **Fabry-Perot Interferometers for Small Displacement Measurements**
L. P. Howard, F. Scire, J. A. Stone (National Institute of Standards and Technology) page 385
5. **A Metrological Three-axis Translator and its Application for Constant Force Profilometry**
V. G. Badami, S. T. Smith, S. R. Patterson (University of North Carolina at Charlotte) page 391

Session VII

Controls and Sensors

Wednesday, November 13, 1996

8:30 a.m. - 10:00 a.m.

Session Chair: Dan E. Luttrell (3M Company)

1. **Modeling and Control of Hysteresis in a Piezoceramic Actuator**
P. Ge, M. Jouaneh (University of Rhode Island) page 396
2. **Implementation of an Acoustic Emission Proximity Detector for Use in Generating Glass Optics**
M. A. Piscotty, J. S. Taylor, K. L. Blaedel (Lawrence Livermore National Laboratory) page 400
3. **Adaptive Pole-zero Cancellation in Grinding Force Control**
H. E. Jenkins, T. R. Kurfess (Georgia Institute of Technology) page 406
4. **Diamond Point Tool Force Measurement in a Production Environment**
J. Schaefer (Texas Instruments, Inc.), C. Arcona (IBM Corporation), T. A. Dow, G. M. Moorefield II (North Carolina State University), K. J. Falter (Taylor Hobson / Pneumo), and R. L. Rhorer (National Institute of Standards and Technology) page 412

Poster Session II

Wednesday, November 13, 1996

10:30 a.m. - 12:00 p.m.

CONTROLS

1. **Submicrometer Overshoot Control of Rapid and Precise Positioning**
S. B. Chang (Precision Instrument Development Center, National Science Council) page 416
2. **Torsional Compliance in Grinding Chatter**
R. D. Entwistle, B. J. Stone (The University of Western Australia) page 420
3. **Closed-loop Control of a Magnetic Servo-Levitated Fast-Tool Servo System**
H. M. Gutierrez, P. I. Ro (North Carolina State University) page 424
4. **Performance Evaluator for Open Architecture NC Boards**
M. Hatakeyama, S. Ueno (Technical Research Institute of Japan Society for the Promotion of Machine Industry) page 428
5. **Fabrication of Surface Perturbations on Inertial Confinement Fusion Targets**
B. H. Jared, T. A. Dow, K. P. Garrard, G. M. Moorefield II (North Carolina State University), and C. Barnes, R. D. Day, D. J. Hatch, L. J. Salzer, G. Rivera (Los Alamos National Laboratory) page 432
6. **Autonomous Cutting Parameter Regulation by Using an Adaptive Modeling and Genetic Algorithms**
T. J. Ko, H. S. Kim (Yeungnam University) page 436
7. **Development of a Thin Brittle Filter Insertion Machine for Optical WDM**
K. Kudo, F. Ohira, K. Koyabu, M. Matsunaga, Y. Inoue (Nippon Telegraph and Telephone Corporation) page 442
8. **Multi-sensor Monitoring on Cold Form Tapping of Large Diameter Internal Threads in High Strength Steel⁺**
Ma, P., Wang, M. (Nanjing University)
9. **Autonomous Acquisition of a Shared, Machining Condition Database in an Open CNC Environment**
M. Mitsuishi, H. Okabe, M. Hashiguchi, K. Tanaka, T. Nagao (The University of Tokyo) page 446
10. **Review of Developments in Sensors for Untended Machining**
A. C. Okafor, A. Ukpong (University of Missouri-Rolla) page 450

⁺Abstract not available at press time.

11. **Integrated Electrodynamic Multicoordinate Drives — Modern Components for Intelligent Motions**
C. Schäffel, E. Saffert, E. Kallenbach (Technical University of Ilmenau) page 456
12. **Design of a Long Range Fast Tool Servo System Using Magnetic Servo Levitation**
B. A. Stancil (Harris Corporation), and P. I. Ro (North Carolina State University) page 462
13. **A Neural Network-based Real-time Compensation Technique for Frequency Output Sensor Arrays**
G. Y. Tian, Z. X. Zhao, R. W. Baines (University of Derby) page 467
14. **Development of NC Program Evaluator for Higher Machining Productivity**
X. Yan, K. Shirase, M. Hirao, T. Yasui (Kanazawa University) page 472

PROCESSING

1. **Local Lapping System to Obtain Wide and Flat Surface with Actively Controlled Force Sensor**
K. Matsumoto, M. Nakao, Y. Hatamura (The University of Tokyo) page 476
2. **Dynamic Friction Polishing of Diamond Utilizing Sliding Wear by Rotating Metal Disc**
K. Suzuki, A. Ide, N. Watanabe (Nippon Institute of Technology), N. Yasunaga (Tokai University),
Y. Seki (De Beers Japan), T. Uematsu (Toyama Prefectural University) page 482
3. **Three-body Abrasive Wear Characteristics Under Reciprocating Motion of CFRP in Vibrating Environment**
S. Teraoka (Kanazawa Technical College), and K. Ishikawa, T. Nakagawa
(Kanazawa Institute of Technology) page 486
4. **Surface Integrity of Polished Infrared Optics and Glass**
V. C. Venkatesh, W. K. Chee, K. H. Tan (Nanyang Technological University), and
K. L. Pey, A. Trigg (Institute of Microelectronics) page 490

DIAMOND TURNING

1. **Burr and Shape Distortion in Micro-grooving of Optical Components**
J-H. Ahn, H-S. Lim (Pusan National University), and D. A. Dornfeld
(University of California at Berkeley) page 496
2. **An Empirical Tool Force Model for Precision Machining**
C. Arcona (IBM Corporation), and T. A. Dow (North Carolina State University) page 500
3. **Effect of Grain Orientation and Grain Boundaries on the Surface Topography of Diamond Turned OFHC Copper**
E. Brinksmeier, W. Preuss, O. Riemer, J. Schmütz (University of Bremen) page 506
4. **Design and Application of In-process Acoustic Sensors for Maximum Sensitivity – Methods and an Example in Diamond Tooling**
J. Daniel, D. A. Dornfeld (University of California at Berkeley) page 510
5. **Development of a Micro-cutting/Testing System Under Scanning Electron Microscope**
M. Mitsuishi, N. Sugita, T. Nagao, Y. Hatamura (The University of Tokyo) page 515
6. **Holographic Testing of Diamond Turned Aspherics. How Well Does it Work?***
W. C. Stenton (Hughes ELCAN Optical Technologies)
7. **Ultraprecision Metal Cutting Process Monitoring Using Acoustic Emission — Burr Formation**
J. Tang, X. Chen, D. A. Dornfeld (University of California at Berkeley) page 520

MATERIALS

1. **Controlled Microcutting Tests on Single Crystal Silicon**
B. V. Tanikella, R. O. Scattergood (North Carolina State University) page 524

*Abstract not available at press time.

NOVEL MANUFACTURING PROCESSES

1. **Patterning Submicrometer Features Using Masked, Broad Beam Ion Machining**
H. E. Caggiano, T. G. Bifano (Boston University) page 532
2. **Nanogrinding**
H. H. Gatzert, J. C. Maetzig (Hanover University) page 536
3. **Effect of Vibration Using Ultrasonic and Low Frequency Vibrations on Grooving Process for Hard and Brittle Materials**
H. Suwabe, K. Ishikawa, M. Uneda (Kanazawa Institute of Technology) page 542
4. **Precision Trimming of Large GFRP Parts Using a Mobile Robot with Onboard Grinder**
Y. Nasu, K. Mitobe, K. Shibata (Yamagata University), N. Kouda (Hitachi Yonezawa Electronics Co., Ltd.), and T. Kuroda (Hitachi Chemical Co., Ltd.) page 546
5. **Thermal Cutting of Plate Glass**
K. Sadashivappa, M. Singaperumal (Indian Institute of Technology), and C. S. Bhaskar Dixit (Bapuji Institute of Engineering and Technology) page 551
6. **Precision Polishing of Optical Fiber Connector**
H. Suzuki, T. Kuriyagawa, K. Shoji (Tohoku University) page 557

GRINDING

1. **In-process Detection of Ductile-brittle Transition in High Speed Scratching of Glass—Theory and Experiment**
J. Daniel, D. A. Dornfeld (University of California at Berkeley) page 561
2. **On-line Measurement of Grinding Wheel Wear Using Acoustic Emission**
W. P. Dong, L. Annecchino, J. A. Webster (University of Connecticut) page 566
3. **Development of Vibration Cutting Using OD-Blade**
K. Ishikawa, H. Suwabe, M. Uneda (Kanazawa Institute of Technology) page 572
4. **Prospective of Fractal Application in Grinding**
X. M. Liu, G. Q. Cai, H. W. Zheng (Northeastern University), D. L. Ma (Liaoning Industrial Art School), and Q. Y. Wen (Shenyang Donghang Mechanical & Electrical Technology Institute) page 578
5. **Grinding Characteristics of Plasma-sprayed Ceramics**
K. Narayanasamy (Indian Institute of Technology), and S. Gowri (Karnataka Regional Engineering College) page 583
6. **The Effect of Workpiece Modulation on Grinding Force**
W. Qu, M. H. Miller (Michigan Technological University) page 589
7. **Development of Grinding Models for Brittle Materials**
K. W. Sharp, R. O. Scattergood (North Carolina State University) page 593
8. **High-efficiency Mirror Finishing Characteristics of Ceramics with Extremely Fine Grit Diamond Wheels**
X-J. Wang (SpeedFam Co., Ltd.), and K. Syoji, T. Kuriyagawa (Tohoku University) page 598
9. **Shear-mode Grinding Force Criteria in Grinding Silicon Wafers with Cup Wheels**
Y. Yoshida (Tokyo Diamond Tool), P. Nan (Northeastern University), K. Imai, H. Hashimoto (Kanagawa Institute of Technology), S. Sekita, I. Iida (Okamoto-Shibayama Kikai), and N. Tsuda (Tokai University) page 602
10. **Cryogenic Grinding of Brittle Materials**
H. Yu, S. Terashima, N. Morita, Y. Yoshida (Chiba University) page 606
11. **The Plastic Zone Formation in Alumina Associated with Ultraprecision Grinding**
I. Zarudi (The University of Sydney) page 610

SLIDEWAY MATERIALS

1. **Friction Characteristics of Linear Plain Bering Guideway and Motion Controllability of Numerically Controlled Slide**
A. Kanai, M. Miyashita, T. Hatai, M. Yoshida (Ashikaga Institute of Technology) page 614
2. **Tribology of Hard Bearing Coatings**
A. P. Santavy, R. O. Scattergood (North Carolina State University) page 618

DESIGN II

1. **Non-Isotropy of Young's Modulus of Elasticity in Very Thin Metal Sheets**
P. V. Pistecky (Delft University of Technology) page 622
2. **High Precision X-Y- θ Micropositioning Stage Using Monolithic Flexure-pivoted Linkages**
J-W. Ryu, D-G. Gweon (KAIST) page 626
3. **Innovations in the Design and Development of Ultraprecision Machines — Lightweight and Ultrastiff Single Point Diamond Turning Machine with Submicron Machining Accuracy**
J. P. M. B. Vermeulen, P. C. J. N. Rosielle, P. H. J. Schellekens (Eindhoven University of Technology) . . . page 631
4. **Straightness Errors of Rectangular Beams Caused by Ambient Air Temperature Gradients***
K. L. Wasson (Aesop, Inc.), and A. H. Slocum, J. H. Lienhard (Massachusetts Institute of Technology) . . page 637

Session VIII

Machine Tool Metrology

Wednesday, November 13, 1996

1:30 p.m. - 3:00 p.m.

Session Chair: Joseph D. Drescher (UTC-Pratt & Whitney)

1. **Using Chebyshev Norms for Estimating Machine Tool Error Parameters (Invited Paper)**
H. Tajbakhsh, Z. Abadin, P. M. Ferreira (University of Illinois at Urbana-Champaign) page 638
2. **A 5-degree-of-freedom Laser System for Rapid Machine Volumetric Error Measurement and Compensation***
Q. Ma, X. Chu, K. Lau (Automated Precision, Inc.) page 642
3. **Laser Measurement Instrument for Fast Calibration of Machine Tools**
P. S. Huang, Y. Li (SUNY Stony Brook) page 644
4. **Practical Thermal Error Correction for CNC Machine Tools**
J. P. Allen, S. R. Postlethwaite, D. G. Ford (The University of Huddersfield) . . . page 648

Session IX

Industrial Metrology

Wednesday, November 13, 1996

3:30 p.m. - 5:00 p.m.

Session Chair: Kirsten M. Carr (Ford Research Laboratory)

1. **Methodologies for the Evaluation of Systematic Form Deviations for Inside-diameter Cylindrical Features and Their Relationship to Process Variables**
R. P. Henke, R. M. Cassou, K. D. Summerhays (University of San Francisco), J. M. Baldwin (Sandia National Laboratories/Livermore), and C. W. Brown (Allied Signal Federal Manufacturing & Technologies) page 654
2. **Best Fit Transformation and Uncertainty Analysis in Computational Metrology**
H-T. Yau (National Chung Cheng University) page 660
3. **Uncertainty of the Coordinate Measurement Process Due to Unwanted Asperities**
J. G. Salsbury (University of North Carolina at Charlotte) page 664
4. **Three-dimensional Metrology Frame for Precision Applications**
M. J. Prather, J. F. Cuttino, D. E. Schinstock (The University of Alabama) page 668
5. **Assessment of Plateau Honed Surface Texture Using Wavelet Transform**
X. Liu, J. Raja, H. Sannareddy (University of North Carolina at Charlotte) page 672

*Extended abstract not available at press time.

Technical Papers



A S P E

POSITIONING PERFORMANCE OF A MAGLEV FINE POSITIONING SYSTEM

John B. Wronosky, Tony G. Smith, Jay D. Jordan, Joel R. Darnold
Sandia National Laboratories, Albuquerque, NM 87185-0501

ABSTRACT

A wafer positioning system was recently developed by Sandia National Laboratories for an Extreme Ultraviolet Lithography (EUVL) research tool¹. The system, which utilizes a magnetically levitated fine stage to provide ultra-precise positioning in all six degrees of freedom, incorporates technological improvements resulting from four years of prototype development experience. System enhancements, implemented on a second generation design for an ARPA National Center for Advanced Information Component Manufacturing (NCAICM) project, introduced active structural control for the levitated structure of the system.

Magnetic levitation (maglev) is emerging as an important technology for wafer positioning systems in advanced lithography applications. The advantages of maglev stem from the absence of physical contact. The resulting lack of friction enables accurate, fast positioning. Maglev systems are mechanically simple, accomplishing full six degree-of-freedom suspension and control with a minimum of moving parts. Power-efficient designs, which reduce the possibility of thermal distortion of the platen, are achievable. Manufacturing throughput will be improved in future systems with the addition of active structural control of the positioning stages.

The demand for smaller critical dimensions in integrated circuits has driven projection lithography to shorter wavelengths. Research and development to extend this trend to extreme ultraviolet (EUV) wavelengths, in the range of 11 nm to 14 nm, is underway at Sandia National Laboratories in Livermore, California. An EUVL laboratory tool using a 10x reduction Schwarzschild camera and magnetically levitated wafer stage driven by a digital feedback controller facilitate this research. Position stability and accuracy must be precisely controlled to achieve circuit overlay of 100 nm features.

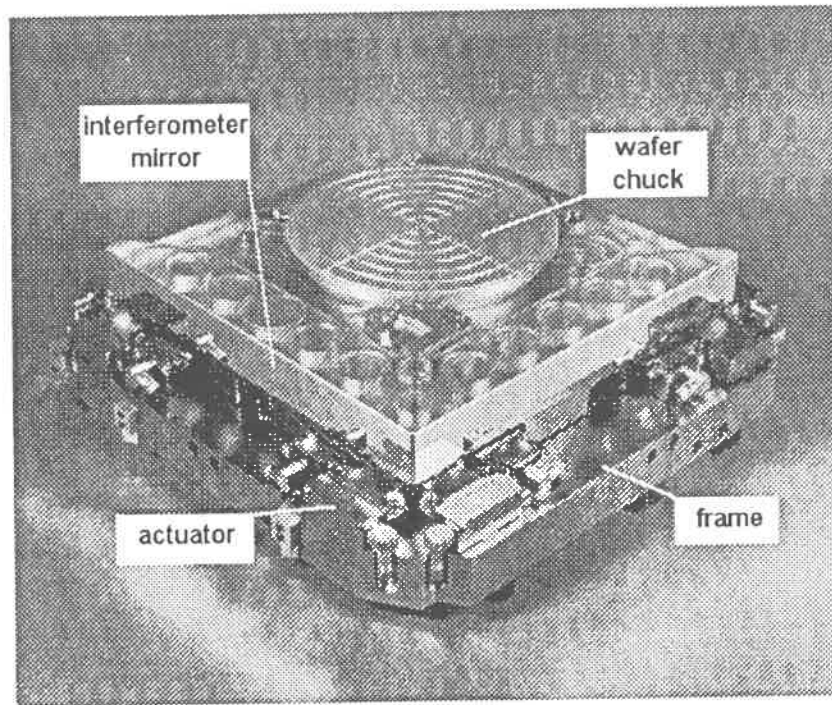
The great expense of the fabs required for integrated circuit manufacturing motivates the emphasis on positioning speed. Increasing fab throughput by increasing the speed of the processing equipment is an economic necessity. For lithography exposure tools, this means minimizing the time spent positioning the wafers between exposures. Unfortunately, regardless of the positioning mechanism, there is a positioning speed limit where further increase in speed results in excitation of structural resonances that degrade accuracy to below acceptable levels. The NCAICM Lithography Structural Control Testbed project develops an active method of structural resonance cancellation. Active structural control is expected to facilitate positioning bandwidth improvement of five to ten times that of the present maglev system.

This paper describes the design, implementation, and functional capability of the maglev fine positioning system. Specifics regarding performance design goals and test results are presented.

¹The EUVL and NCAICM programs, including the subject positioning systems, being performed at Sandia National Laboratories, is supported by ARPA DALP and by the US Department of Energy under contract DE-AC04-94AL85000.

The Maglev Stage

In the late 1980s, a concept for a magnetically levitated fine positioning technique was developed at MIT by Dr. David Trumper. Refinements of the technique, and the design of a control system for its implementation, have resulted in the fine positioning stage shown in the figure below.



The stage consists of the levitated fine stage platen containing sixteen ferrous targets and the interferometer mirror, the frame holding the sixteen E-core electromagnets and six capacitive position sensors. A Hewlett Packard laser interferometry system consisting of five interferometers, two laser sources, beam benders, optical receivers, and high resolution laser axis boards provide position information in all six degrees of freedom. The interferometry can resolve to 0.618 nm in X and Y and 1.236 nm in Z (used only on the EUVL system). The Z axis requires that a wafer be in place to provide a reflecting surface for the light beam. The maglev fine stage utilizes six capacitive sensors for determination of the "gap" between the movable platen and stationary frame. This information is used to linearize the maglev actuator force characteristics for control purposes. The fine stage can be controlled and positioned using only these sensors for feedback, but with less accuracy than that obtainable with the interferometers.

Control System Electronics

Two computers are used in the implementation of the wafer positioning system. An embedded VME-based 486 PC provides the user interface, and a four-processor TMS320C40 board furnishes the control computations. The electronics include a total of seven TMS320C40 DSPs for data acquisition, manipulation, and control. Analog-to-digital I/O is used to gather gap information from the capacitive sensors as well as other miscellaneous system data. Sixteen

channels of digital-to-analog I/O control the current amplifiers which drive the electromagnetic actuators.

Custom current amplifiers are used to supply drive current to the sixteen electromagnetic actuators. They are 1 amp limited, but capable of up to 100 volts output. This provides the inductive magnetic actuator with a fast force response thereby allowing high bandwidth positioning.

System Performance

Measures of system performance are accuracy, repeatability, stability, time required to move to a requested position, and power consumption. For this system, accuracy and repeatability are determined by the performance of the laser interferometry system and the interferometer mirror. The specifications given by Hewlett Packard for the interferometer depend on the application, but both accuracy and long-term repeatability can both be a small fraction of a μm when atmospheric disturbances are not present (as in the vacuum environment seen on EUVL). Errors introduced by mirror distortions are minimized by virtue of the mirror's kinematic mounting scheme, the good thermal properties of the Zerodur mirror material, and the low level of heat dissipated in the maglev actuators.

Elements necessary for achieving very stable positioning are quiet sensors and actuators and a relatively vibration-free environment. The interferometers satisfy the requirement for low noise, as they have less than one least-significant bit of short-term noise. The electromagnetic actuators are also very quiet, due in part to the fact that the force provided is a function of the current in the actuator and the actuator inductance is large. The vibration-free environment for these systems is generally provided by setting the system on a table that is isolated by pneumatics or elastomers from the floor vibrations. This attenuates high-frequency motion, but amplifies motion at the resonant frequency of the isolation system, which the maglev control system must reject. In order for the maglev system to be able to compensate for this motion, the bandwidth of the control system must be higher than the resonant frequency of the isolation system. Another source of vibrations that can cause significant difficulty at this level of accuracy is sound waves. The amplification and transmission of ambient acoustics by the isolated table can easily dominate the system error. Degradation of stability by a factor of four has been observed when the system is exposed to the noise in a clean room without an acoustic enclosure around the isolated table. In quiet surroundings, positioning stability of 1.5 nm RMS (one axis) has been achieved. For two axes, the actual position of the platen is within 5 nm of the desired position 99.3% of the time.

The time required to complete a move is a function of many variables, including the force available from the actuators, the dispatch with which the force can be produced, the bandwidth of the control system, the maximum velocity capability of the interferometers, and the rigidity of the supporting structure. The actuators are capable of producing force sufficient to generate accelerations of over 1g. The rapid generation of the required force is hindered by the large inductance of the actuators, so 100V amplifiers are used to effect the desired actuator currents in a timely manner. High controller bandwidth is crucial to quick step and settle times, but can be quite difficult to achieve due to system structural resonances. These resonances cause the control system to become unstable as the bandwidth is increased. To overcome this problem, piezoelectric actuators are embedded in the levitated platen and are used to suppress the

resonances induced by the maglev controller. A bandwidth of 150 Hz was attained in this manner. The laser interferometer system is capable of a peak velocity of 400 mm/sec. This is not a limiting factor on short moves, as that velocity may only be required for a brief duration or not at all. The primary problem that must be overcome to enable rapid step-and-settle times is the lack of rigidity in the supporting structure. This includes the table the system is mounted to, the structure holding the interferometry components, and the isolation system. Most of these items are not difficult to address, but their importance must be understood by the equipment designers - these are not generally elements that can be fixed with a band-aid late in the development cycle. The biggest challenge is the isolation system, because of conflicting requirements placed upon it. The isolation system must be compliant to provide the desired isolation from floor vibrations, but when a large move is undertaken this compliance allows the impulse seen by the isolation system to set the assembly into motion at its resonant frequency. There are several potential fixes, including the use of an active isolation system, shaping the move profile to minimize residual motion, and modifying the isolation system to pass the reaction forces directly to a mechanical ground. An active isolation system has been tried, but with limited success. The other approaches have not been tried, and would be appropriate avenues for future work.

To illustrate the capabilities of the system as it now exists, a move of 20 mm (a reasonable step size for semiconductor photolithography) is executed in a time-optimal fashion with peak available acceleration of 10 m/sec^2 and peak attainable velocity of 400 mm/sec. The move is completed in under 90 msec, with the deviation from the desired trajectory never exceeding $15 \mu\text{m}$. Due to the residual vibrations, however, it takes the system another 70 msec for the error to drop below $1 \mu\text{m}$, and another 100 msec beyond that to reach steady state. Again, this situation could be greatly improved by the application of the techniques noted above.

The last measure of performance mentioned above is the power dissipated in the actuators. This is important to minimize any thermal distortions that could degrade the positioning accuracy. By using permanent magnets to offset gravity, the power dissipation is made very low. Power consumption is on the order of 0.1W while stationary, and only a few watts are dissipated during a move, depending on the size of the move, the bandwidth of the controller, and the maximum acceleration requested.

REFERENCES

1. Tichenor, G. D. Kubiak, S. J. Haney, R. P. Nissen, K. W. Berger, R. W. Arling, A. K. Ray-Chaudhuri, K. B. Nguyen, R. H. Stulen, J. B. Wronosky, J. D. Jordan, T. G. Smith, J. R. Darnold, P. M. Kahle, A. A. Jojola, S. M. Kohler, R. S. Urenda, D. R. Wheeler, J. E. Bjorkholm, O. R. Wood, II, G. N. Taylor, and R. S. Hutton, "**Recent results in the development of an integrated EUVL laboratory tool**," Proc. of the SPIE Symp. on Electron-Beam, X-Ray, and Ion-Beam Submicrometer Lithographies for Manufacturing V, vol. 2437, 292 (1995).
2. R. W. Arling, S. M. Kohler: "**Six Degree of Freedom Fine Motion Positioning Stage Based on Magnetic Levitation**", Second Annual Symposium on Magnetic Suspension Technology, Seattle, WA., August 11-13, 1993.
3. J. B. Wronosky, T. G. Smith, J. R. Darnold: "**Development of a Wafer Positioning System for the Sandia Extreme Ultraviolet Lithography Tool**", Third Annual Symposium on Magnetic Suspension Technology, Tallahassee, Florida, December 13-15, 1995.
4. J. B. Wronosky, T. G. Smith, J. R. Darnold, J. D. Jordan,: "**Maglev six degree-of-freedom fine position stage control system**", The Signal Processing Applications Conference & Exhibition at DSPX 1996.

Design and Characterisation of Nanometer Precision Mechanisms

(Paper I)

Ying Xu, Paul D Atherton, Malachy McConnell and Thomas R. Hicks
Queensgate Instruments Ltd, Waterside Park, Bracknell, RG12 1RB, UK

1. Introduction

In recent years, as the result of rapid developments in various fields of precision engineering there has been a big increase in the need for precision positioning and scanning systems capable of nanometer or, sometimes, even subnanometer resolution and repeatability. This trend is expected to grow, requiring new design concepts and techniques for exploration of more novel devices to meet the demands of various applications. For example, wafer steppers are making silicon chips with line widths down to 200nm. Scanning Probe Microscopes are used to establish how well such chips are made, and the introduction of M-R head technology allows 1Gigabyte disks to become the norm, while 23 Gigabyte disk stacks are now available for desk top PCs. These machines, and the machines which make these machines, combine advanced optical design with advanced motion control technology, which can position components to an accuracy of a nanometer or below.

To meet the challenge of these developments, Queensgate Instruments is developing an ultra-precision positioning technology which combines Queensgate's piezoelectric and NanoSensor technology into multi-axis positioners with the ability to position with sub-nanometeric accuracy. This system is composed of a series of stages, called NanoMechanisms, including single axis stages, x-y stages, tilting stages and yaw stages etc.. The combinations of these stages can provide three, four, five or six degree of freedom positioning units.

2. Design Philosophy of Nanometer Precision Mechanisms (NanoMechanisms)

Only piezo-electric devices have the potential to meet the resolution requirements with very high stiffness. However, an external sensor is required to control their position because piezo-electric devices are non-linear and exhibit hysteresis. The capacitance micrometer is ideally suited to this task, being small and simple with an intrinsic resolution capability which is effectively infinite. To achieve a pure single axis motion, a flexure guiding mechanism is used, which implements constraints to any off axis motions and combines piezo actuator and sensor together to form an integral stage system. Figure 1 is a typical closed loop control block diagram of this kind of system. In the diagram, the motion measured by the sensor is fed back to the controller which moves the stage to minimise the difference between the sensed motion and the command. In this case, the positioning precision in the metrology loop is mainly determined by the capabilities of sensor and controller.

Capacitance position sensor

The capacitance sensor developed by Queensgate is called the NanoSensor. This is a highly linear device with linearity errors of $< 0.02\%$ over specified operating range (normally between $100 \sim 500 \mu m$). Operating over a reduced range it can produce even higher linearity. Super Linearity $< 0.01\%$ is achieved by using an NXB-L sensor (which has a larger sensing area) over the $100 \mu m$ range (its normal working range is $500 \mu m$). The NanoSensor has a positional noise level in normal operation of $< 0.005 nmHz^{1/2}$ (RMS) and can be fabricated from very stable materials like SuperInvar or Zerodur. It is non contact, and is free from hysteresis. It also has the advantages of being very compact, simple, cheap, and with no power dissipation at the point of measurement. It is thus well suited to the measurement of very small displacements.

Controller

When designing a system with 0.1 nm resolution and 100 μm range the ability to access that range under computer control is normally extremely difficult, since it is a dynamic range of 1 part in a million, or 20 bits. To address this problem Queensgate have developed a Digital Signal Processor based control system, which has an intrinsic resolution of more than 21 bits, and is addressable digitally. It should be noted that this far exceeds the resolution of most A/D and D/A converters. Advanced digital algorithms can reduce the effect of mechanical resonances improving settle times greatly. The working bandwidth can be controlled by computer and loop parameters be defined by the user for performance optimisation. Using such a controller it is possible to measure the non-linearity and to compensate for it. Furthermore once parasitic motions such as roll pitch or yaw in such mechanisms have been characterised it is possible to compensate for them in complex multi-axis system. The linearity errors can be fully compensated to < 0.01 %. Below that the measurements are normally limited by the intrinsic linearity of the calibration systems.

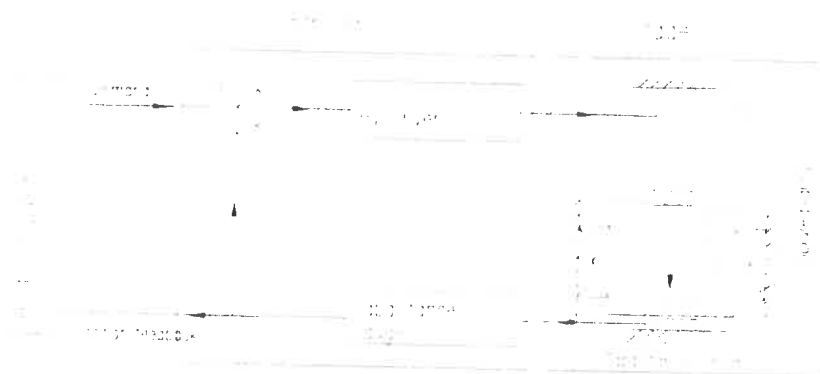


Fig. 1. Block diagram of a single axis control system of NanoMechanism

The use of this compensating technique is very important to achieve a metrological capability at the nanometer level. It is obvious that within the control loop the sensor is not absolutely linear, so the system linearity can be further improved by software compensation. Ideally, the mechanism should exhibit pure orthogonal motions - that is, an x-y device must only have degrees of freedom along the x and y axes. In reality, there exist uncontrolled (parasitic) motions arising from distortion due to internal forces and to manufacturing limitations. The errors from these parasitic motions have been minimised by optimising the mechanical design and can be reduced further by the compensation technique. Note that the errors can only be compensated if they are predictable, i.e. the parasitic motions have to be, not only measurable, but also repeatable.

3. Design concepts and considerations

Coordinate system

First it is necessary to define the coordinates used to describe positions. The obvious system to use for positioning stages is an orthogonal Cartesian Coordinate system, as shown in figure 2. With this one can define a position with its X,Y,Z coordinates and an arbitrary rotation as components of rotation about the X,Y and Z axes. More usefully one can describe a movement as a change in the X,Y and Z coordinates. As drawn, the positive Z direction is up out of the page and the negative Z direction down into it. Rotations are described with respect to the X,Y and Z axes in a right-handed sense, thus $+\theta$ is a clockwise rotation when looking along the +Y direction, $+\phi$ is a clockwise rotation when looking along the +X direction and $+\gamma$ is a clockwise rotation when looking along the +Z direction. They are normally called pitch, roll and yaw in aeronautical terms.

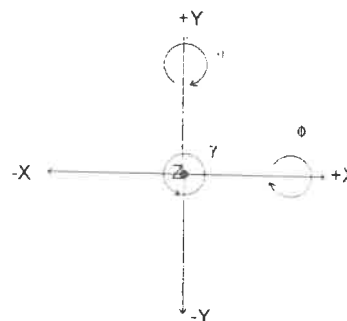


Fig.2 Coordinate System

Positioning accuracy: the concept of trueness

To move the stage a position command is sent to the controller by a computer. The motion is produced by a piezo actuator and monitored by a sensor. Using the feedback signal, the controller moves the stage to minimise the difference between the sensed motion and the command. How small the difference can be is mainly determined by the controlling ability of the system and can be interpreted as how precisely the stage can be positioned. It is obvious that the positioning precision will be affected by resolution, noise level and reproducibility (drift and hysteresis) of the system. Moreover, if the stage motion is measured with an external measuring device which is assumed to be a perfect system there will be a difference between the commanded position and desired position, and how close they are is defined as mapping trueness. Therefore, the final positioning accuracy should be determined by both the positioning precision and the mapping trueness. How these are dealt with in NanoMechanism design will be discussed later.

Resolution and noise

The resolution of measurement or positioning is directly related to the noise level of the system. A peak to peak noise level is not easily measured or interpreted, since with any noise distribution the longer you wait the larger the deviation from the mean. Therefore the rms value is normally used which can be measured with standard equipment. The noise amplitude distribution is important when looking at resolution. Usually Gaussian noise dominates and in this case the rms is equivalent to the standard deviation, sigma. 67% of samples taken will be within one sigma of the mean value. That means there is a 67 % chance of resolving two features which are a distance of two sigma of the noise apart, as shown in figure 3, (or 90 % chance of resolving two features which are six sigma apart).

The noise power spectrum is a most important piece of information, since it can show up the underlying sources of noise - such as mains pick up, which is localised at 50 or 60 Hz. Figure 4 demonstrates a measurement of the noise power spectrum of the DSP based NPS3000 controller built for NanoMechanisms, showing a noise level of $< 10 \text{ pmHz}^{-1.2}$. In the test, the NPS3000 controller is used to control a single axis stage, NPS-Z-15B, in closed loop with a stage working bandwidth of 100 Hz. The noise signals plotted are from the HV drive voltage applied to the piezo actuator. 50 Hz mains pick up can be seen clearly, although at a very low level.

The noise in the NanoMechanism system is, in general, composed of sensor noise, piezo drive noise and mechanical and acoustic noise. Sensor noise will be interpreted by the control loop as a command, thus becoming actual displacement noise. The feedback signal from the sensor is used to generate a drive voltage to be applied to the piezo actuators. Noise will be introduced in this process and contribute to the stage positioning noise. The effect of this noise can be measured by the sensor and, therefore, at least partially servoed out. The ability of the system to servo out the drive noise depends on the bandwidth set: the higher the bandwidth the better the contribution is servoed out. External mechanical inputs such as ground vibration and acoustic noise will also cause the stage to move. The effects of these

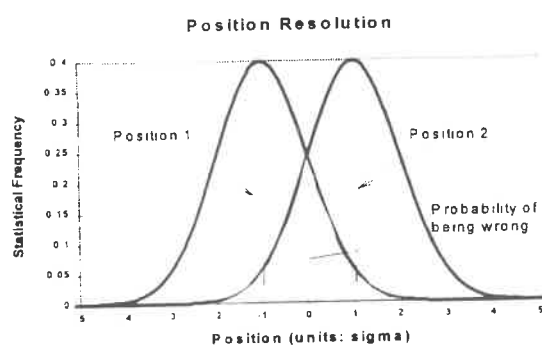


Fig. 3. Resolving two positions

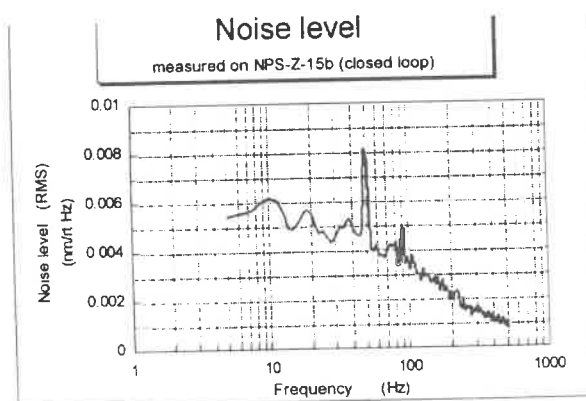


Fig. 4. Noise spectrum of NGC3000 controller

inputs can be minimised by increasing the stiffness of the stage. It can also be servoed out if the system bandwidth is sufficiently high. For the control system NPS3000, the measurement bandwidth can be set up to 12 kHz and closed loop bandwidth 2 kHz.

Linearity and mapping

In order to apply linearity compensation, and thus to know what the actual motion or position of the stage is, the system has to be calibrated against an external measuring device of higher order accuracy. The command position, x_c , which is the position measured by the internal sensor and the actual position, x_a , can, to a certain extent, be related with a mapping function expressed as $x_c = f(x_a)$. The simple form of the mapping function is a power series

$$x_c = a_0 + a_1 x_a + a_2 x_a^2 + a_3 x_a^3 + a_4 x_a^4 + \dots \quad (1)$$

Ideally $a_0, a_2, a_3, a_4, \dots$ would be zero and a_1 unity. Then the sensor scale factor, a_1 , is the linear factor describing the relationship between the actual stage position as measured by a hypothetical perfectly accurate position sensor and the position measurement fed back to the user's computer. The mapping trueness error is the set of errors on the individual 'a' coefficients, which is then the sensor scale factor error. When the mapping function is first order (a straight line), then the mapping error becomes the linearity error. As an example, a linearity error of 0.05% in a 100 μm range device results in a 50nm absolute position uncertainty between the 0 μm position and the 100 μm position, as shown in figure 5(a). Usually the deviation from linearity is roughly parabolic and in some systems this is easy to compensate for electronically. The result of compensating one, slightly imperfect, parabola with another is usually an S curve of much lower amplitude so the mapping error is much lower, as shown in figure 5(b). This is equivalent to using the a_1 and a_2 terms of equation 1. If one were to use the higher order terms, an even better result could be achieved. This can be done easily in microprocessor based sensor systems or externally in the user's computer. It has been found that there is little to gain in going higher than fourth order, see figure 5(c).

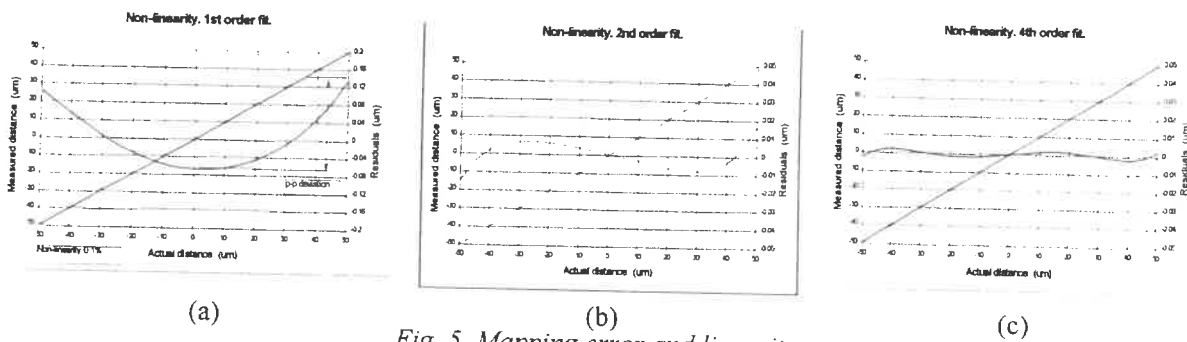


Fig. 5. Mapping error and linearity

Parasitic motions and errors

Parasitic motions in stages can be identified as either angular: yaw, pitch and roll; and linear : out of plane motion, non orthogonality and crosstalk; and will introduce unexpected positioning errors. The parasitic motions caused by distortions can be minimised by careful design and the use of finite element modelling (FEM) and optimisation. In the constrained axes the stiffness should be designed to be as high as possible, and the other axes as low as possible. This is achieved in the NanoMechanisms by properly arranging the flexures and choosing the parameters. However it is, sometimes, limited by the system resonant frequency which because of mode coupling, requires the stiffness in all directions to be high. FEM is able to predict local and global distortions and then the structure can be properly optimised to decouple the forces or make the inevitable distortions cancel each other. As mentioned before, the compensating technique can be used if these parasitic motions are predictable. Note that these motions are a function of the stage position, but are not necessarily linear, leading to complex topologies. The hysteresis may make the prediction very

hard if not impossible. For this reason, the force change in the system is desired to be highly linear and repeatable. Friction may generate a force loop due to the change of its directions and, therefore, is always a source of hysteresis.

When specimens are mounted on to the NanoMechanism, Abbe errors have to be considered carefully due to the parasitic angular motions. Small angular errors can have a large affect at the nanometer level: for example, a tilt of just $1 \mu rd$ with an offset of 1 mm gives a 1 nm position error. To reduce this effect, specimens should be positioned as close as possible to the measuring axes of the sensors. For example, in an x-y-z 3 axis NanoMechanism system the specimen holder is located at the point which is co-linear with the sensor measuring axes, as shown in figure 6. The effects of yaw, pitch and roll of the x-y stage can thus be minimised.

The linear parasitic motions such as non-orthogonality or crosstalk are mainly affected by manufacturing tolerances. The axes of two pairs of sensors in the x-y stage have to be very orthogonal to each other and co-linear to the moving axes of the platform. Using modern manufacturing technology the deviation of the orthogonality of the sensor axes can be generally controlled within 0.5 mrd which gives an orthogonality error of $0.5nm/\mu m$ (i.e. 0.05%) in the x-y plane.

By kinematic mounting, the position reference becomes traceable and the distortions from thermal expansion and driving force can be decoupled. This is important for the stage to have metrological capability at the nanometer level. Even for a SuperInvar stage of the size of 100×100 mm, $1^\circ C$ temperature change will cause 30 nm change in dimension ($\alpha = 0.3 ppm/^\circ C$), and the frame bending from the driving force is normally in the range of tens to hundreds of nm. Otherwise a position uncertainty of about the same magnitude might be introduced into the system.

Dynamic characteristics

Aside from the metrology and accuracy of motion, the dynamic performance of the system is also important because the stability and speed are critical to many applications. Ideally there would be no phase lag between command and position, and the mechanism would respond perfectly to a step input - no rise time, over shoot, or settle time.

For a linear, second-order, damping-free mechanical system, the resonant frequency is determined by system stiffness and mass. In an optimally designed mechanism, the stiffness is usually dominated by the stiffness of piezo stacks in its translation axis. For a stage with motion amplification, the effective stiffness of the piezo actuator will be reduced as $k_e = k_p/Q^2$, where k_p is the stiffness of piezo and Q is the amplification. Reducing the mass can increase the system resonant frequency. However, as the mass of the platform decreases the stage performance becomes more sensitive to the influence of the load mass, i.e. the resonant frequency will drop down rapidly as the mass of the specimen increases. Dynamic properties of the system can also be improved through other approaches such as applying a proper damping or using advanced servo control techniques. In instrumentation, the design specifications often use the criterion of settling time, defined as the time required for the system to settle within a certain percentage of the input. For NanoMechanisms, like other instruments, settling time is a more direct description of the dynamic performance than resonant frequency. For a piezo driven NanoMechanism, the settling time consists of the slewing time and the time taken for the resonant oscillations. The former is dominated by the slew rate which is determined by the capacitance of the piezo stacks and the current drive capability of the drive electronics. For second order systems the requirement typically specifies a maximum delay before the output remains within 2 % of its final value after a step input change, which takes a duration of four time constants ($4\tau = 4/\xi\omega_n$).² From this it can be seen that the system response can be improved by increasing both the resonant frequency and the damping factor. Normally, flexure bearing stages are highly resonant with very low damping factors. Therefore, extra damping will be very helpful and can

Fig. 6. 3D NanoMechanism

effectively reduce the decay time, but only if it can be introduced without friction, which can cause hysteresis. A typical step response of a NanoMechanism is shown in figure 7, which is measured from the single axis stage NPS-Z-15B. It demonstrates that the settling time of the stage for a $3\text{ }\mu\text{m}$ step is less than 5 msec (settling to within 2 % of the step).

5. Some examples of NanoMechanism devices

NPS-Z-15A/B: It is a single axis linear motion stage which is designed to produce a pure motion along the z axis. The stage has a closed loop range of $15\text{ }\mu\text{m}$ and a linearity of $< 0.06\%$ (without compensation) with subnanometer resolution. A tube shaped compact flexure mechanism is designed into the stage to decouple the parasitic off axis and tip-tilt motions from piezo stacks. The pitch is measured to be less than $1\text{ }\mu\text{rd}$ over whole range, (without flexure mechanism the pitch can be as high as about $15\text{ }\mu\text{rd}$).

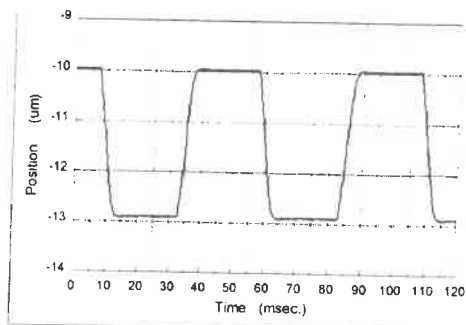


Fig. 7. Step response of NPS-Z-15b

NPS-XY-100: This is a two axis linear motion stage with a $\Phi 40$ mm aperture in the middle and thus convenient for scanning applications. It has a dynamic range of $100 \times 100\text{ }\mu\text{m}$ with subnanometer resolution. With a yaw compensating mechanism in the stage, the yaw error is easily controlled to be less than 2 arcsec. and roll and pitch are insignificantly small. Integral kinematic mounting mechanism helps to release the strains induced by internal drive forces and temperature changes, improving the stability of the system. Combining NPS-XY-100 and NPS-Z-15 forms a 3D positioning and scanning system, as shown in figure 6, ideal for metrological SPM applications.

NPS- $\Theta\Phi$ Z-xx: The tip-tilt-z stage has a closed loop range of $6\text{ mrd} \times 6\text{ mrd} \times 25\text{ }\mu\text{m}$ with a resonant frequency over 2 kHz. With this stage, linear motion with subnanometer resolution and angular motion with nano-radian resolution are achievable.

6. Conclusions

Some of the technologies used in Queensgate's nanometer precision mechanisms have been introduced and discussed to explain how a metrological capability at nanometer or even subnanometer level can be achieved. This involves both optimised design of the system and advanced compensating techniques. More detailed information is available from Queensgate.³ A series of NanoMechanisms, ranging from single axis to multi axis stages, have been designed and built. The combination of these stages can provide motions of up to six degrees of freedom with nanometer precision. The design considerations discussed here also deal with the possible parasitic motions which may introduce positioning errors and uncertainty to the system. Some approaches for avoiding or reducing these errors were mentioned. A comprehensive assessment of metrological characteristics is a complicated and long term project, especially for multi axis systems. Initial testing has shown promising results, such as low noise level, very small parasitic motions, high linearities and good step response. Further results will be reported in the near future.

Reference

1. P. D. Atherton, Y. Xu and M. McConnell, "New x-y stage for positioning and scanning", *Proceedings of SPIE's, Annual Meeting*, Aug. 1996, Denver, USA
2. S. T. Smith and D. G. Chetwynd, *Foundations of Ultraprecision Mechanism Design*, Gordon and Breach Science Publishers, 1992
3. *The NanoBook*, Queensgate Instruemnts Ltd, 1996

A PIEZO-DRIVEN TEM STAGE WITH SUBNANOMETER POSITION STABILITY

H. van der Wulp, H.F. van Beek, and B.M. Mertens[†]

Delft University of Technology

Faculty of Mechanical Engineering, Laboratory for Micro Engineering

[†]Department of Applied Physics, Particle Optics Group

Delft, The Netherlands

Introduction

Since their invention in 1931, Transmission Electron Microscopes (TEMs) have become very important "optical" instruments because they can offer atomic resolution (~ 0.1 nm). TEMs are applied in solid-state chemistry and physics, material science, and biology, where they are used for (fundamental) research and for the development of new products and techniques. In order to get atomic resolution, the position stability of the specimen relative to the objective lens is very important.

In most TEMs, the specimen is placed at the end of a long specimen holder which is beared in the goniometer [1]. By means of the goniometer, which is placed on the outside of the TEM column, the specimen is positioned in the center of the vacuum chamber in between both objective pole pieces.

The necessary high position stability can be achieved by referring the specimen position directly to the lower pole piece of the objective lens. Then, the actual position of the specimen must be adjustable relative to the point of reference. Depending on the experiment, positioning in one, or perhaps as many as five Degrees of Freedom (DoF) is needed, implying translations of several millimeters and rotations up to 80° . The overall specimen position repeatability should be better than 10 nanometer. Furthermore, the specimen drift rate (quasi-static) should be lower than 0.1 nm/min., while the specimen vibration amplitude should be smaller than 20 pm. Both these levels of specimen stability should be obtained within 10 seconds after positioning [2].

Apart from the application of an improved specimen holder, a very direct solution for the realization of a stable image would be the use, in a feedback loop, of the momentary image position information obtained from the TEM CCD-camera. Unfortunately, this method has a severe practical handicap: the bandwidth of this CCD-image feedback loop is limited to tenths of a Hertz as a result of the necessary complex cross-correlation calculations between images. This means that image feedback can only be used for the compensation of the quasi-static changes in specimen position.

So, for the improvement of the specimen position repeatability and stability, a new specimen holder concept has been developed. The new specimen holder is called 'nano-stage'; the nano-stage table carries the specimen. The stage is meant to be completely incorporated in between the objective lens pole pieces of a (future) TEM. Furthermore, the nano-stage must properly operate in ultrahigh vacuum and in the presence of an electron beam and strong magnetic fields.

Experimental Approach

The nano-stage has to be remarkably small so that it can be placed inside the vacuum chamber, onto the lower pole piece of the objective lens, as can be seen in Figure 1. Miniaturization has two advantages: it increases the eigenfrequency and therefore the intrinsic stability of the nano-stage relative to the pole piece and it makes structural mechanical loops as short as possible.

The main problem of the new concept is the combination of the required specimen motion in five DoF over large strokes, within the limited space of the vacuum chamber.

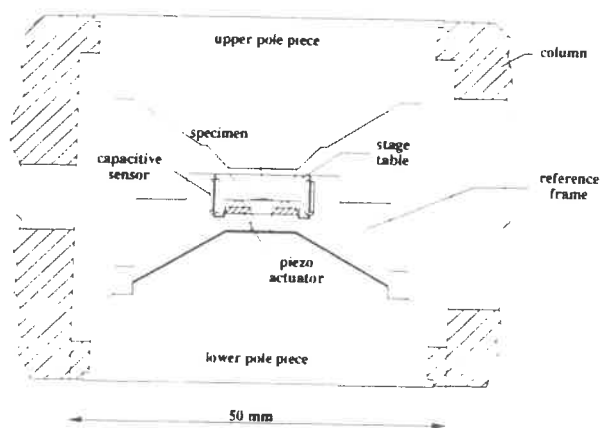


Fig. 1. The future TEM nano-stage, carrying the specimen, must be very small and at the same time allow for five DoF motion over large strokes, while being mounted on the lower pole piece of the TEMs objective lens.

At the start of this project, piezo actuators were chosen to move the nano-stage table. These actuators are small, have picometer positioning resolution, possess a large bandwidth and do not create a magnetic field. However, the deformation of small piezo actuators is limited to several hundreds of nanometers, depending on the piezo type and its dimensions. Besides, they suffer from non-linearity, hysteresis and creep. To overcome the drawback of the limited piezo deformation, the principle of Inertial Sliding Motion (ISM) [3] is used, a method often applied in Scanning Probe Microscopes (SPM). It enables the realization of long strokes.

A drawback of ISM is the variation in step size. This variation amounts to about 10% of the nominal step size (i.e. 100 nm).

Consequently, in order to obtain nanometer stage table position repeatability, accurate position sensors are needed to determine the momentary stage table position. Here, capacitive position sensors were chosen because of their picometer resolution, vacuum compatibility and extremely low energy loss. By means of a reference frame, the sensors use the lower pole piece as the mechanical position reference. The application of the alternating current bridge method for the determination of capacitances, in combination with a phase sensitive detection method [4], resulted in a range-resolution ratio of 10^6 at 20 Hz bandwidth.

Because the capacitive sensors are used to measure position changes with the highest resolution possible, only quasi-static (here <20 Hz) position changes can be compensated for. Therefore, the nano-stage should be isolated from dynamic vibrations, in order to obtain high stage table stability and thus high image resolution. In practice, the TEM column is already vibration isolated by a specially designed mass-spring-damper system. By placing the nano-stage into the vacuum chamber, the nano-stage is much better isolated from both acoustic vibrations and thermal changes resulting from external heat sources than the current TEM specimen holder. Furthermore, due to the miniaturization of the nano-stage, the eigenfrequency is enhanced considerably and, consequently, the isolation from mechanical vibrations is strongly increased. The same applies for the nano-stage intrinsic stability.

Subsequently, two different compensation concepts are available to minimize (quasi-static, sub-nanometer) specimen drift: feedback by means of position sensors with picometer resolution and control on the basis of knowledge about the drift characteristics of the stage table. Application of the first concept asks for a combination of two sets of capacitive sensors: one set to obtain nanometer position repeatability over a millimeter range and one set to obtain drift measurements with picometer resolution for drift compensation. The main disadvantage of this concept is the

space needed to implement an extra set of position sensors. This paper concentrates on the position control based on knowledge of the stage table drift characteristics.

Measuring set ups

Two of the nano-stage set ups used during this research, were described earlier [5,6]. They are shown in Figures 2a and 2b. The difference between these set ups is their size and the kind of piezo actuators used: piezo tubes and shear piezos. To verify the stage table stability, the capacitive sensor resolution of the set ups is 15 and 30 pm, respectively, at 20 Hz bandwidth. The set ups were placed in a temperature conditioned room outside the TEM, on a vibration isolated table which was placed under a perspex box to minimize air flow disturbance.

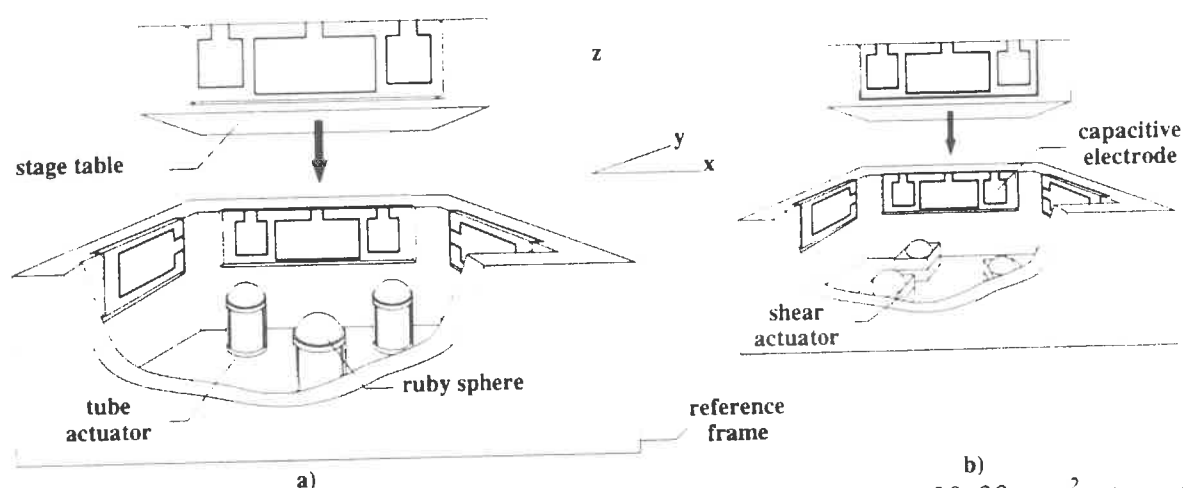


Fig. 2. Two nano-stage set ups for positioning in the xy -plane. a) A $30 \times 30 \text{ mm}^2$ stage table driven by piezo tubes. b) A $20 \times 20 \text{ mm}^2$ stage table driven by shear piezos.

A third set up was built according to our ideas with regard to a future TEM nano-stage (see Figure 1) in order to investigate the nano-stage table stability by verifying the image stability. The stage table is $14 \times 14 \text{ mm}^2$ and supported by three ruby spheres, which are on top of single shear piezos. The stainless steel stage table is equipped with a V-shaped groove. As a result, the TEM nano-stage has only one DoF with 0.1 mm range. The shear piezos are attached to an aluminum reference frame, which is on its turn attached to the pole piece, that was, for this purpose, slightly modified. The stage table carries an amorphous carbon film. The stage table position is measured by a set of capacitive position sensors, having 30 pm resolution at 20 Hz bandwidth. The 50 nm thick carbon specimen is illuminated by an electron beam of about 500 nm diameter ($I \cong 1 \text{ nA}$). The specimen image can be recorded by a CCD-camera with a maximum recording rate of 1 Hz. The CCD-camera has 512×512 pixels. Converted to the specimen scale this means 0.2 nm per pixel. By means of pixel interpolation, image shift measurement may be improved to subpixel level.

Intrinsic nano-stage stability

Figure 3 relates to the stage shown in Figure 2a. It shows the frequency spectrum from DC up to 3.2 kHz and a zoom-in from DC up to 0.2 Hz of the analogue position signal in the Y-direction of the capacitive phase sensitive detector at 2.2 kHz sensor bandwidth. The nano-stage first eigenfrequency is between 1.5 and 2 kHz. The table is in rest. From Figure 3 can be concluded

that due to miniaturization and vibration isolation, the stage table position changes are only quasi-static. The intrinsic stage table stability in the Y-direction can be given by a DC drift rate: < 30 pm/min. The intrinsic stage table stability in the Z-direction is even better: < 20 pm/min.

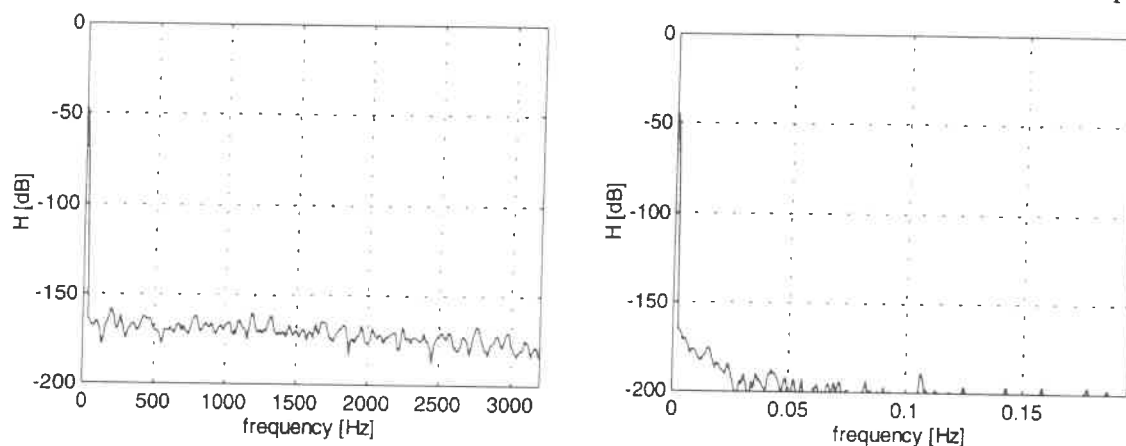


Fig. 3. Frequency spectrum of the capacitive sensor output signal, representing the table position in the Y-direction when the stage table is in rest. a) Spectrum from DC up to 3.2 kHz. b) A zoom-in from DC up to 0.2 Hz.

Stability after ISM

The stage table position stability directly after ISM is strongly influenced by hysteresis and creep of the piezo actuators, and by creep in the epoxy layers and in the contact points between stage table and ruby spheres.

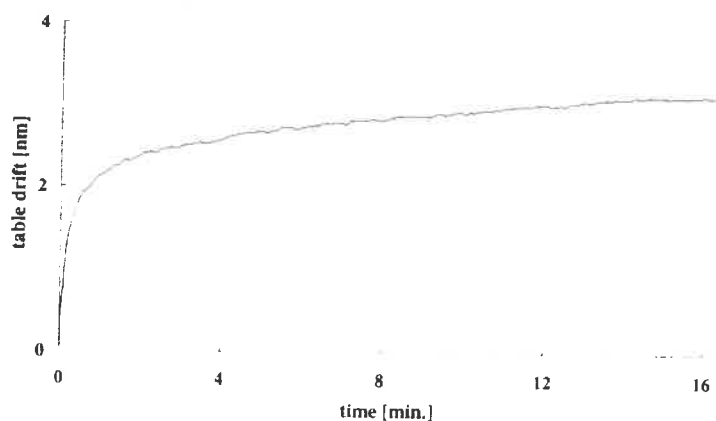


Fig. 5. Stage table drift after a 120 V, 300 nm ISM step measured with the 'piezo tube' nano-stage sketched in Figure 2a.

The stage table drift increases with increasing piezo voltage and is as small as 1 % of the step size. The amount of drift, of course, depends on the piezo deformation history and on the sign of the driving voltage. The vertical lines give an impression of the drift repeatability.

The stage table drift curves after a single ISM step can be well predicted by a logarithmic curve. This, and the reasonable repeatability of the drift curve makes drift compensation possible based on the knowledge of the stage table drift characteristics, resulting in drift and drift rate decrease.

However, for higher driving voltages, variation in the amount of drift becomes large, making this type of position control less effective. At best, drift compensation can be avoided by using only small driving voltages. For single 20 V, 35 nm ISM steps, stage table drift rates are generally lower than 0.1 nm/min. within a few seconds after the ISM step is made!

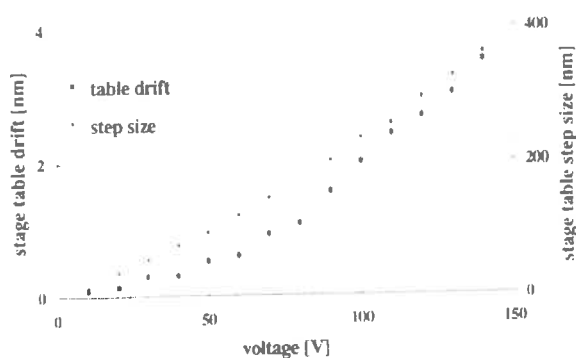


Fig. 6. The stage table step size and the stage table drift value 3 min. after a single ISM step vs. the piezo voltage, measured on the 'piezo tube' stage.

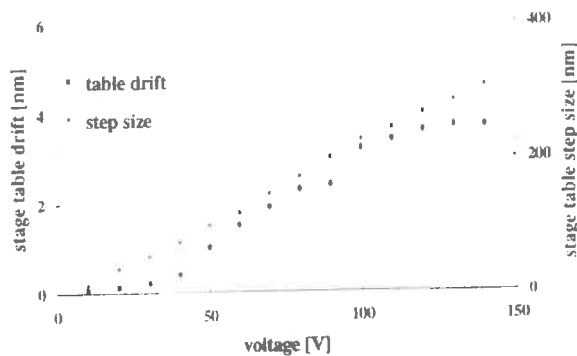


Fig. 7. The stage table step size and the stage table drift value 3 min. after 10 ISM steps vs. the piezo voltage, measured on the 'piezo tube' stage.

Another advantage of small driving voltages is a low heat production. The only disadvantage of using small ISM steps is a low stage table velocity. The nano-stage of Figure 2b gives similar results.

In practice, multiple ISM steps are applied to move over large distances. The drift after multiple ISM steps does not differ much from that after a single step, as can be seen in Figure 7. It shows stage table drift after 10 successive ISM steps for different driving voltages. Again, the drift rate at 20 V driving voltage is lower than 0.1 nm/min. within a few seconds after stepping. Increasing the number of steps to 50 and 100, respectively, had no significant influence on the amount of stage table drift. Only a change of stepping direction temporarily increases drift to the 50 V ISM step level. An alternative and more simple option to minimize stage table drift after ISM positioning is to compensate a single or multiple of ISM steps by a single ISM step in the opposite direction. The size of the last, opposite step depends on the number and size of the previous ISM steps.

The TEM nano-stage

Figure 8 shows the effect of switching on the electron beam, on the nano-stage table drift recorded by the capacitive sensors on the one hand and on the image drift recorded by the CCD-camera on the other hand. As a result of specimen heat up, a huge image drift occurs. During the first 60 seconds of the measurement, image drift is large than 100 nm so that no correlation can be made between the first two images. The stage table position, however, is not affected at all. The same effect occurs when the stage table is moved to a new position over a distance with the same order of magnitude as the electron beam diameter. In that case, a new thermal equilibrium must develop. This means that the image stability is not just dependent on the stage table stability, but also on the internal specimen stability. It is obvious that this heat effect decreases with decreasing travelling distance.

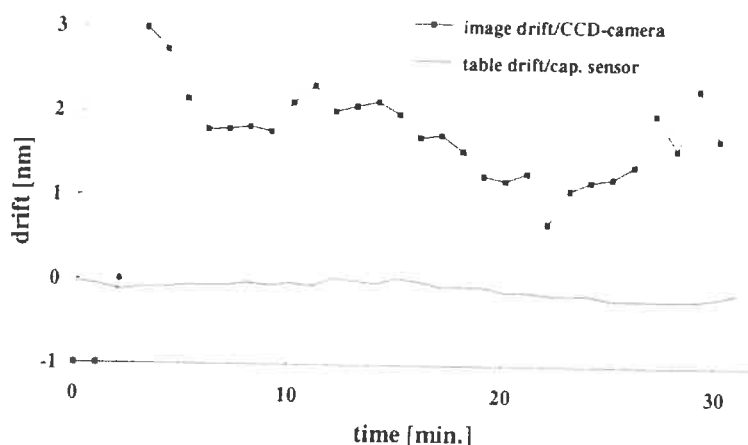


Fig. 8. Image drift and stage table drift after switching on the electron beam ($I \cong nA$).

At the same time, Figure 8 shows a second problem. After thermal equilibrium is obtained, the image stability seems to be low compared to the stage table stability. However, this low image stability may also result from the relatively low image resolution and from errors in the pixel interpolation calculations due to this low image resolution and due to the low image contrast. Therefore, the

image resolution and the image contrast should be improved to be able to compare stage table drift and image drift at the subnanometer level.

Conclusions

With the nano-stages described, nanometer position repeatability over a stroke as much as one millimeter and tens of picometer position stability, a few seconds after ISM steps, can be achieved. Due to miniaturization and vibration isolation, stage table position changes directly after positioning are only quasi-static. Stage table drift is mainly caused by piezo creep, which is clearly related to the applied piezo voltage. Because of its logarithmic decrease with time, stage table drift, which is as small as 1 % of the ISM step size, can be well predicted, making drift compensation possible. However, for small ISM steps, stage table drift can be neglected, making drift compensation superfluous.

The image stability in a TEM is not just dependent on the stage table stability, but also on the internal specimen stability. This internal specimen stability is mainly depending on thermal and perhaps on other material effects within the specimen.

Acknowledgments

The support of Marcel Hoogteijling, Peter Pisteccky, and Pieter Kruit is gratefully appreciated. This work was supported by the Foundation on Fundamental Research on Matter (FOM), Utrecht, The Netherlands.

Literature

- [1] P.E.S.J. Asselbergs et al., "Design of a new, five-axis motorized goniometer for the CM series microscopes: the CompuStage", Philips Electron Optics Bulletin, 1990.
- [2] P.E.S.J. Asselbergs, Philips Electron Optics, private communication.
- [3] D.W. Pohl, Rev. Sci.Instrum. **58**, 54 (1987).
- [4] W.Chr. Heerens, J.Phys.E. **19**, 897 (1986).
- [5] H. van der Wulp and R. van Dalen, Proc. 8th Intern. Prec. Eng. Seminar, Compiègne, 1995.
- [6] H. van der Wulp and M. Hoogteijling, Proc. 5th Intern. Conf. on New Act., Bremen, 1996.

ULTRA-PRECISE TRAVELING SLIDER FOR FABRICATION OF DVD GLASS MASTER

J. Nishida, N.Kikuri, N.Uchida, and N.Takasu
Toshiba Corporation
Kawasaki, Japan

1. Introduction

With the advent of DVD(Digital Video Disc), the memory capacity of optical disks is being increased to cope with moving pictures. The capacity of DVD is 9.4 G bytes/disk. The diameter is 120 mm, track-pitch 0.74 μm , and the allowable error of track-pitch is ± 30 nm. The authors have developed an ultra-precise traveling slider for a mastering machine to fabricate DVD glass masters. The slider consists of not only a coarse-motion table, but also a fine-motion table, to realize the precise traveling. The fine-motion table is controlled to compensate for the positioning error of the coarse-motion table. In this paper, mechanical construction, the method to compensate for the error, and evaluation of the positioning error of the developed slider are described.

2. Mastering machine

2.1 Outline of mastering machine

Figure 1 shows the schematic view of the mastering machine. The exposure laser beam which is modulated by AOM(Acousto-Optic Modulator) is led to the objective lens. The objective lens is moved vertically by VCM(Voice Coil Motor) so that the exposure laser beam focuses just on the glass master. The slider moves the objective lens across the glass master precisely at the designed velocity, so that the line of spiral pits is recorded on the glass master.

2.2 Specifications of the slider

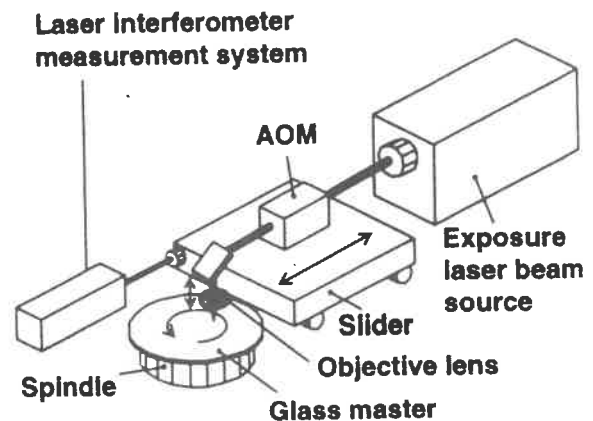


Figure 1 Schematic view of the mastering machine

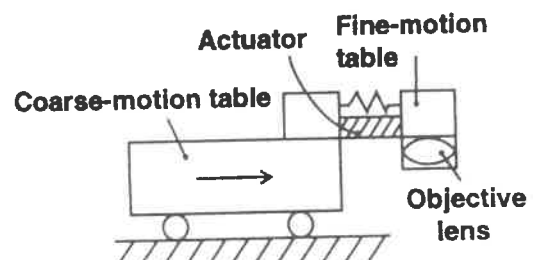


Figure 2 Principle of the compensation control method

The slider moves at low velocity of about 10 $\mu\text{m/s}$ for 1 or 2 hours during recording. The designed position of the slider is given as the function of time. The positioning deviation from the designed position (traveling deviation) is the main cause of track-pitch error. The traveling deviation of the slider is required to be within ± 15 nm (5σ) to maintain the track-pitch error within ± 30 nm. It is difficult to maintain the deviation within ± 15 nm for hours using the slider driven by a ball screw, linear motor or friction drive mechanism.

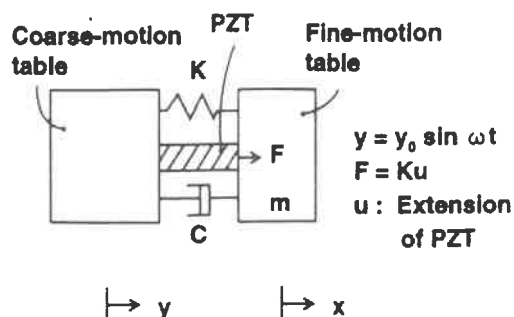


Figure 3 Analytical model of the slider

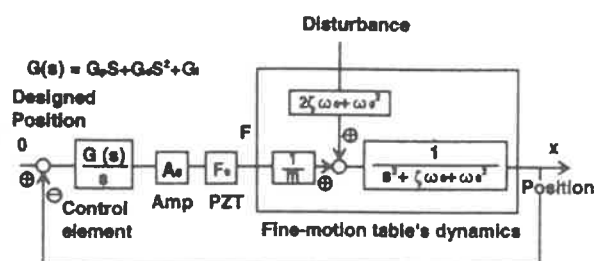


Figure 4 Block diagram of the compensation control method

Therefore, the compensation control method using the fine-motion table is adopted.

3. Compensation control method

3.1 The principle

Figure 2 shows the principle of the compensation control method. In this method, the slider consists of not only a coarse-motion table(CMT), but also a fine-motion table(FMT). The FMT, to which the objective lens is attached, is mounted on the CMT. The positions of both tables are measured by laser interferometer measurement systems, and fed back to the control systems of both tables, which are completely independent of each other. The FMT is controlled to compensate the traveling deviation of the CMT, so that the objective lens travels exactly at the designed velocity. The same designed position is applied to the control systems simultaneously. The CMT moves

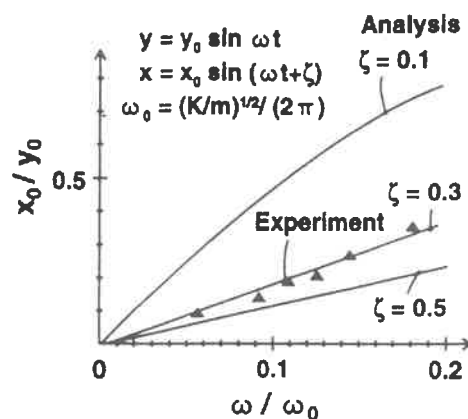


Figure 5 The effect of the compensation control method

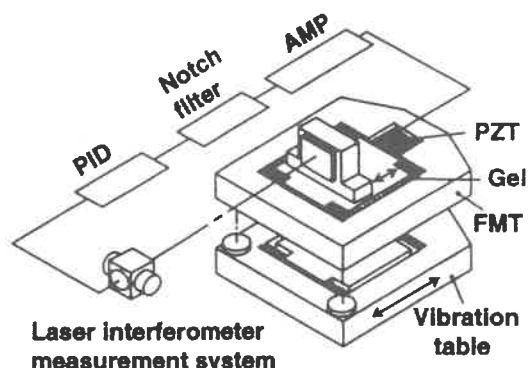


Figure 6 Setup for the experiment of the compensation control method

independently of the FMT, and the FMT moves within the maximum traveling deviation of the CMT. Therefore, the stroke of FMT can become small.

3.2 Effect of compensation

Figure 3 shows the analytical model of the compensation control of FMT. The velocity of FMT is so slow that only the vibration element is considered. The compensation control of FMT is expressed in the block diagram of the control system shown in Fig. 4. If the control method is PIDaction, the transfer function $H(s) (= x/y)$

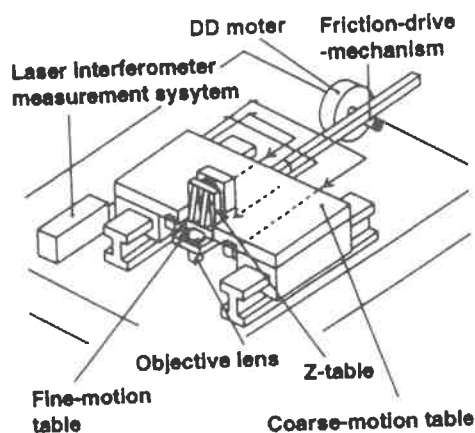


Figure 7 Schematic view of the Slider

Table 1 Specifications of the two tables composing the slider

	CMT	FMT
Stroke	160 mm	10 μ m
traveling deviation	± 50 nm	± 15 nm
Actuator	DD motor and friction-drive mechanism	Air bearings with porous ceramic nozzles
Guided way	PZT	Leaf springs

of this system is as follows;

$$H(s) = \frac{2\zeta\omega_0 s^2 + \omega_0^2 s}{s^3 + (2\zeta\omega_0 + \alpha g_d) s^2 + (\omega_0^2 + \alpha g_p) s + \alpha g_i} \quad [1]$$

ω_0 : FMT natural angular frequency

ζ : FMT damping coefficient

g_p, g_i, g_d : gain of PID action

$\alpha = A_0 \cdot F_0/m$

Supposing that $g_d = 0$, $g_p \ll 1$, the ratio of the vibration amplitude (x_0/y_0) is shown as Eq. (2), if g_i in the stability limit is substituted to Eq.(1).

$$\begin{aligned} \frac{x_0}{y_0} &= |H(j\omega)| \\ &= \frac{(\omega/\omega_0)^2}{1 - (\omega/\omega_0)^2} \sqrt{\frac{1 + 4\zeta^2(\omega/\omega_0)^2}{(\omega/\omega_0)^2 + 4\zeta^2}} \quad [2] \end{aligned}$$

In the case of $\zeta = 0.1, 0.3$ and 0.5 , the ratios of the vibration amplitude are shown in Fig. 5.

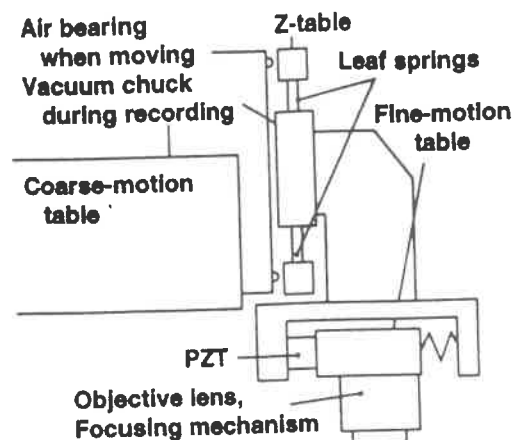


Figure 8 Z-table locked to the Coarse-motion table

The vibration amplitude of FMT is proportional to disturbance frequency. If the frequency ratio (ω/ω_0) is small enough, the vibration amplitude of FMT is reduced.

A simple experiment was conducted as shown in Fig. 6. The FMT, which is supported by leaf springs and driven by PZT (stroke 10 μ m), is oscillated by the vibration table. The FMT is controlled to compensate the vibration using the laser interferometer system (resolution 2.5 nm, sampling frequency 8 kHz). The experimental results are shown in Fig. 5. If the natural frequency of the FMT is designed to be 20 times as high as the disturbed one, the vibration could be reduced to one-tenth.

4. Ultra-precise traveling slider

4.1 Outline of mechanism

Figure 7 shows the schematic view of the developed slider. The slider has a Z-table on the CMT, and the FMT is mounted on the Z-table. The Z-table coarsely moves the objective lens vertically, because the thickness tolerance of glass masters is larger than the stroke of the focusing mechanism. Table 1 shows the specifications of the CMT and the FMT. The CMT, guided by air bearings with porous

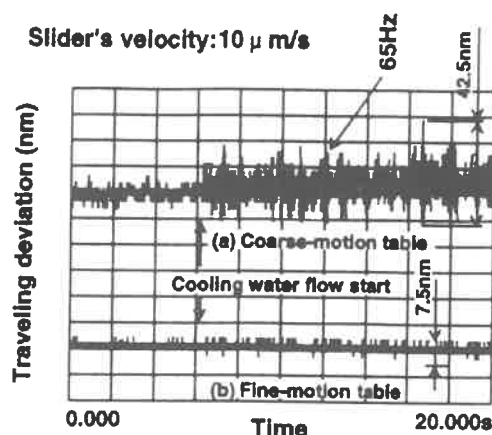


Figure 9 Traveling deviation of the slider

ceramic nozzles, is driven by a DD-motor and a friction-drive-mechanism. The FMT is the same table as used in chapter 3. To compensate for the traveling deviation sufficiently, the Z-table should be guided with high rigidity because the FMT is mounted on it. The Z-table is guided by air bearings when moving, and locked to the CMT by vacuum as shown in figure 8 during recording. The natural frequency of the FMT coupled on the Z-table is 450 Hz.

4.2 Evaluation of the traveling deviation

Figure 9 shows the traveling deviation of the slider. Without disturbance, the CMT travels maintaining the traveling deviation within 12.5 nm_{p-p} at the velocity of 10 μm/s. But with the disturbance caused by the cooling water of the exposure laser beam source, the vibration (65 Hz) occurs and the traveling deviation increases to 42.5 nm_{p-p}. Compensating for the traveling deviation using FMT, the deviation is reduced within 7.5 nm_{p-p} even with the disturbance. Figure 10 shows the deviation during the recording on the glass master at the velocity of 21 μm/s to 6 μm/s. Because of the movement of the focusing mechanism, the deviation is larger than that shown in Fig. 9. But the objective lens could travel through 40

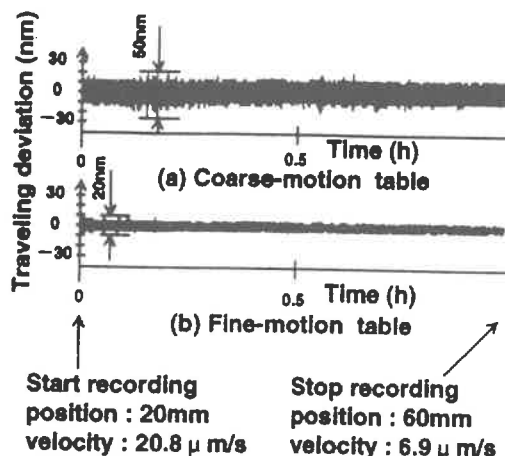


Figure 10 Traveling deviation during recording

mm maintaining the traveling deviation within 20 nm_{p-p} for 1 hour.

5. Conclusion

An ultra-precise traveling slider has been developed for the mastering machine for DVD. The slider consists of not only a coarse-motion table, but also a fine-motion table. The fine-motion table is controlled to compensate for the traveling deviation of the coarse-motion table. Using this system, the following results were obtained.

- (1) Using the coarse-motion table, the traveling deviation was within 15 nm_{p-p} without disturbance, but over 40 nm_{p-p} with disturbance caused by the cooling water in the exposure laser beam source.
- (2) Compensating for the traveling deviation with the fine-motion table, the deviation of over 40 nm_{p-p} is reduced to 7.5 nm_{p-p}.

REFERENCES

- [1]T.Iida et al.,1993,"High Recording Density Optical Disc Mastering", Jap. J. Appl. Phys., Vol. 32, pp. 5447
- [2]S.Sakuta et al.,1996, " Ultra-Precision Positioning by Friction Drive", Proc. of Japan USA Symposium on Flexible Automation, Vol.1,pp.357

AN ELECTRODYNAMIC MICROACTUATOR FOR THE ALIGNING OF AN OPTICAL FIBRE WITH AN INTEGRATED POSITION SENSOR

Thomas Frank

Technical University of Ilmenau, Faculty of Mechanical Engineering

Department of Microsystem Technology

PO-Box 0565, 98684 Ilmenau Germany

Abstract

This paper describes the design, fabrication and testing of an electrodynamic microactuator for moving an optical fibre. This device with two degrees of freedom has a motion range of 200 μm . It is used for the alignment of optical fibres in the sub-micrometer range. The positioning of a monomode optical fibre with respect to one or several other optical fibres or an optical chip is possible. As a result of the electrodynamic principle the motion is free of slide and friction. The technologies of wet chemical etching of silicon and galvanic batch-processes are used. The magnetic field is generated by a permanent magnet. Around the magnet surface copper coils are built for the deflection of the magnet in two dimensions.

Key words: linearmotor, electrodynamic, optics, microcoil, scanning, microactuator, electroplating, positioning

Introduction

The development of microfabricated motors has become an expanding field in microtechnology. The efforts have concentrated on the use of electrical forces mostly with electrostatic, ultrasonic or electromagnetic and electrodynamic actuators^{1,2,3,4,5}. With the new high energy permanent magnetic materials it is possible to create a new generation of electromagnetic and electrodynamic actuators.

microactuator only the forces required are generated and little energy is used.

This paper describes the design, fabrication and testing of an electrodynamic microactuator for moving an optical fibre in two dimensions. The generated forces have an effect on the end of the optical fibre and bend it, see figure 1.

This application is used for the alignment of an optical fibre in the sub-micrometer range. The positioning of a monomode optical fibre with respect to one or several other optical fibres or an optical chip is possible. A kind of special actuator can be used for a fibre-switch for monomode fibres or when creating a topographic picture it is possible to use the actuator for scanning a surface in two dimensions.

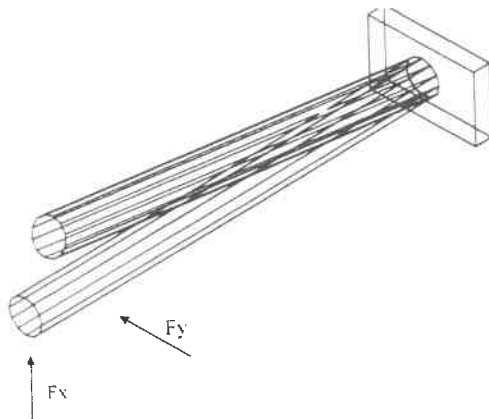


Figure 1 The bending of the optical fibre

One application is the positioning of an optical fibre. For the moving of an optical fibre in a micrometer range forces only in a micro Newton range are needed. Macroactuators generate forces in Newton range and use much energy, that is more than you need for the moving of an optical fibre. The application of a microactuator is possible. With a

Design of the actuator

The actuator is made of two (100)-oriented silicon wafers each structured on both sides. The assembly of the actuator is shown in figure 2. Between the wafer 1 and 2 you can see the permanent magnet, the current paths for the x-deflection and the driven optical fibre. At the upper side of wafer 1 and the bottom of wafer 2 there are the contact pads for the x-deflection current paths and the current paths for the y-deflection. The magnet is fixed to the optical fibre and supported by the optical fibre. Planar copper coils are placed for the two-dimensional deflection of the optical fibre around the magnet. The current lines are embedded in the silicon and they have a width of 100 μm with a spacing of 50 μm and a thickness of 80 μm . The number of windings is 8. The maximum current

density is 500 A/mm². When the actuator is in action, the needed maximum current is 200 mA. At the end of the actuator is the sensor for detecting the position of the moving optical fibre with respect to the other optical fibre or an optical chip.

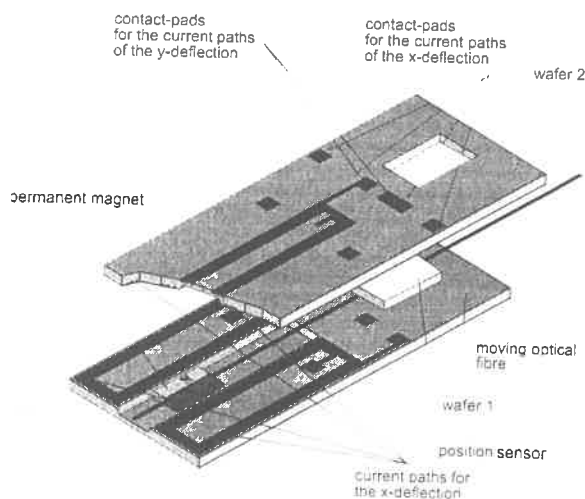


Figure 2 Outline of the actuator

Concept of the microactuator

The motion principle of the actuator is the force on moving electrical charges in a magnetic field (Lorentz force). The force is perpendicular to the current and the magnetic field, according to:

$$\mathbf{F} = I \int (d\mathbf{l} \times \mathbf{B}) \quad (1)$$

For the independent controlling of the two moving directions, separate control forces for both coordinate directions are needed.

Figure 3 illustrates the principle of the forces generation. The current paths are fixed and the magnet is the moving part of the actuator. In figures 3a and 3b the principles of the x- and y-deflections are explained and in figure 3c the combination of parts 1 and 2 is shown. The position of the current paths for the x-deflection is directly above the north and south poles of the magnet. In this area the y-component of the magnetic field predominates. The result is, that these current path generates forces in x-direction (figure 3a). Figure 3b depicts the position of the current paths for the y-deflection. In this area the x-component of the magnet field predominates. The result is, that these current paths generate forces in y-direction. Th. Frank ^{4,6} has recently described more details for the correct position of the current paths.

In the moving range of the actuator the magnet field in the area of the current paths alters when the magnet moves. In this case the symmetrical construction of the actuator compensates the displacement cross-talk of both moving directions^{4,6}.

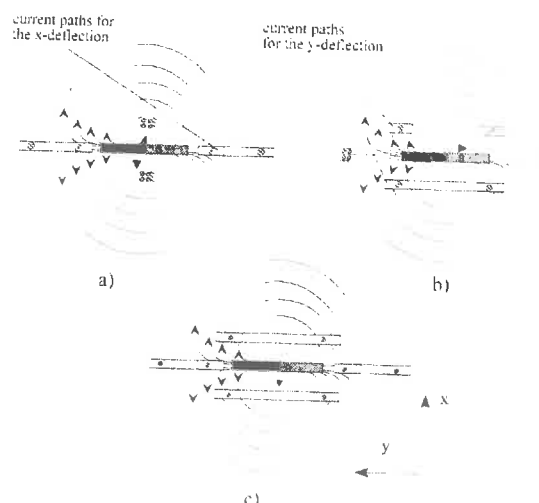


Figure 3 The principle of the force generation, a) Part 1: Principle of the x-deflection, b) Part 2: Principle of the y-deflection, c) Principle of the x/y-deflection with the combination of part 1 and 2

The generated forces have an effect on the optical fibre and bent it. Figure 4 shows the bending of the optical fibre as a result of the generated forces.

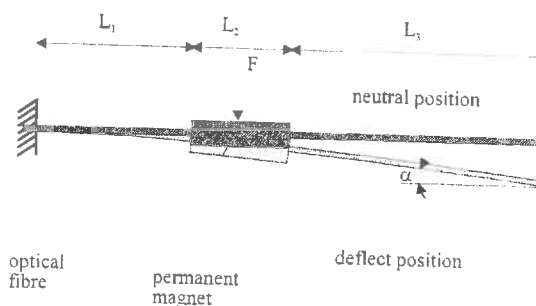


Figure 4 The bending of the optical fibre

Th. Frank ^{4,6} describes the calculation of the generated magnetic field characteristics and the generated forces. The calculated generated forces of the actuator are in the amounts of 5 mN ^{4,6}.

The position sensor is needed for detecting the position of the two connected optical fibres. The sensor measures the intensity of the light in the cladding

which is not coupled into the core but into the cladding. The actuator scans the connected optical fibre or optical chip. The correct position is the position with a minimum of intensity in the cladding, from this is concluded that the intensity in the core is a maximum⁷.

Calculation of the dynamic behaviour

You can describe the dynamic behaviour of the actuator with a spring mass oscillator with two degrees of freedom, hence the system has two resonance frequencies. Figure 5 shows the simple model⁸. The stiffness c_1 represents the stiffness of the optical fibre from the part L_1 (Figure 4), c_2 the stiffness of the part L_2 . The reduced mass of the optical fibre from the part 1 and the mass of the magnet are depicted in m_1 , the reduced mass of the optical fibre from part 2 in m_2 .

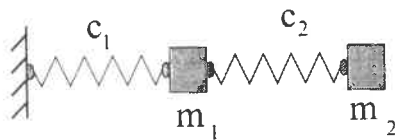


Figure 5 Dynamic model of the actuator

By the variation of the magnet position a variation of the two resonance frequencies is possible. Figure 6 shows the two resonance frequencies versus the position of the magnet L_1 , the total length is constant.

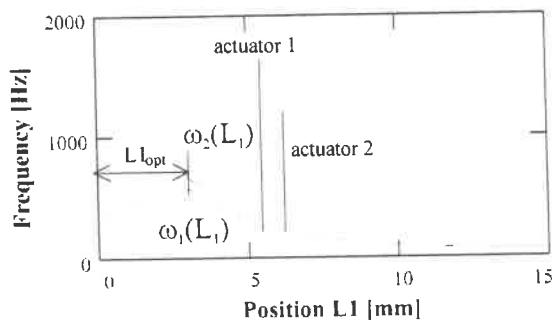


Figure 6 The resonance frequency versus the position of the magnet L_1 . L_{1_opt} is the optimal position of the magnet.

For a position drive a high resonance frequency is advantageous. In Figure 6 the optimal position of the magnet is the maximum from ω_1 and the minimum from ω_2 . In Figure 6 you see the two resonance frequencies of the two realized actuators. They have not reached the optimum. The optimal length of the optical fibre (part L_{11}) is very short, therefore, this part is too stiff and the actuator has not enough power to

implement the moving demands. In figure 9 and 10 you can see the two resonance frequencies.

Fabrication process

For the fabrication two 3" (100)-oriented silicon wafers were used. Figure 7 shows the separate steps. In the first step the three-dimensional structure for the installation of the optical fibres, the moving space for the magnet and the channels for the copper coils were made by an anisotropic chemical etching process. A thermal oxide was used for the masking film. All structures were etched until the natural etch stop at the (111) planes of the silicon was reached (step 1).

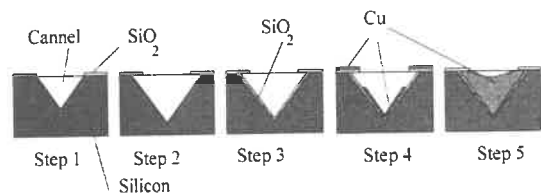


Figure 7 Separate steps for the copper coils production. Step 1: Anisotropic chemical etching, step 2: Isotropic chemical etching, step 3: Second thermal oxidizing, step 4: Metallized with chrome and copper, step 5: Electroplating

After the anisotropic etching, the (111) planes were etched by an isotropic etching process in (111) direction (step 2). Step 2 produced a freely standing oxide masking film.

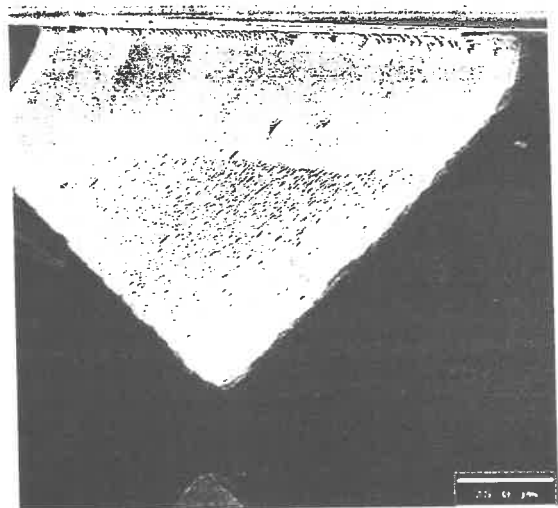


Figure 8 Scanning electron micrograph of the structure after metallized with chrome and copper (step 4)

Immediately after these processes they were cleaned and another thermal oxidizing treatment followed. The oxide film insulates electrically the copper coils and the silicon wafer. After the oxidizing the wafers were metallized with chrome and copper.

In figure 8 the result is shown. The first oxide masking film overshadows a part of the structure. In the overshadow range it is not metallized and it is insulated from the wafer-top. In the following electroplating process the structure for the current paths was only contacted, with the result that the copper was grown only in the structure, see figure 9. Without the overshadowed part the copper would grow at the complete wafer top. At the end the two wafers were mounted together by epoxy resin.



Figure 9 Scanning electron micrograph of the structure after electroplating (step 5)

Experimental results

The results of the very accurate measurement with a laser-vibrometer are shown in figures 10, 11 and 12. Figure 10 shows the position of the fibre end versus the current in the deflection coils for the y-direction. The current in the deflection coils for the x-direction is constant. The three different currents deflect the fibre end in the following x-positions $x = 0 \mu\text{m}$, $100 \mu\text{m}$, $-100 \mu\text{m}$.

The gradient of the current-position curve for the y-part of the actuator increases in the case when the actuator moves in y-direction and has an additional movement in x-direction. This means, that the y-deflection with an additional movement in x-direction is higher than the y-deflection in the neutral

position, when the current is constant. The reason is the non-linear increase of the magnetic induction with reducing distance r from the magnet. The magnetic induction is increased proportional with $1/r$ near the magnet⁹. For the x-part of the actuator the same applies.

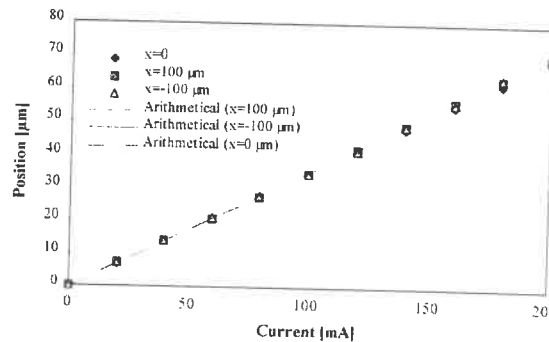


Figure 10 The position of the magnet fibre versus the current in the deflection coils for the y-direction. The current in the deflection coils for the x-direction is constant.

Figure 11 shows the frequency response of the actuator 1 with the deep resonance frequency. The frequency response for the y-direction is the same.

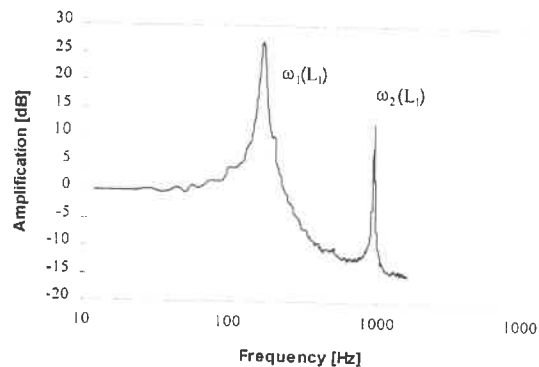


Figure 11 Amplification versus the frequency of the optical fibre (actuator 1)

Figure 12 shows the result of shortening the optical fibre at the part L_1 . The first resonance frequency ω_1 rises and the second resonance frequency ω_2 falls. The displacement cross-talk between the x- and y-direction is smaller than 2% in the motion range of $\pm 100 \mu\text{m}$ in both directions^{4,6}.

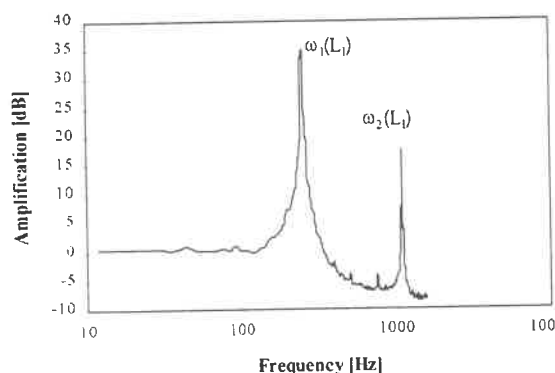


Figure 12 Amplification versus the frequency of the optical fibre (actuator 2)

Conclusions

The presented microactuator was built up in silicon by the use of microstructuring technologies. The accessible exactness is in the sub-micrometer range. Forces higher than 5 mN and a frequency of up to 100 Hz can be realized. A result of the symmetrical design is a minimum of displacement cross-talk. The embedded current paths tolerate a high current, because of the large cross section.

Acknowledgement

The presented work is financially supported by the German Federal Ministry of Education, Science, Research, and Technology (contract number: 13MV0316) as well as the State Ministry of Science, Research, and Culture of Thuringia, Germany (contract number: W6-918/3-183).

References

- 1 B.Wagner, M. Kreutzer, W. Benecke, "Permanent Magnet Micromotor on Silicon Substrates", *Journal of Microelectromechanical Systems*, Vol.2, No. 1, pp. 23-29, March 1993
- 2 "Aktoren aus mikrostrukturiertem Silizium für pneumatische und hydraulische Abwendungen", Herausgeber: VDI/VDE-Technologiezentrum Informationstechnik GmbH
- 3 S. Bütgenbach, "Mikromechanik", B.G. Teubner Stuttgart 1991

- 4 T. Frank, Realisation of a Microactuator for the two-dimensional Alignment of Optical Fibres, pp 136-140, *Actuators 96*, 5th International Conference of New Actuators, 26- 28 June, Bremen, Germany
- 5 Chong Ahn, Yong j. Kim Mark G. Allen, "A Planar Variable Reluctance Magnetic Micromotor with Fully Integrated Stator and Coils", *Journal of Microelectromechanical Systems*, Vol.2, No. 4, pp. 165-173, December 1993
- 6 Th. Frank, Development and manufacturing of a two-dimensional microactuator for move optical fiber, SPIE's Micromachining and Microfabrication '96 Symposium, 14-15 October 1996, Austin TX
- 7 T. Gerlach, D. Enke, Th. Frank, L. Hutschenreuter, H-J. Schacht, R. Schüler, Towards an integrated microsystem for the automatic adjustment of mono-mode optical waveguides, MME'96, Barcelona, 21-22 October 1996
- 8 K.Magnus, "Schwingungen", B.G. Teubner Stuttgart 1986
- 9 E. Philippow, "Taschenbuch der Elektrotechnik", Band 1 Grundlagen, VEB Verlag Technik Berlin 1968

Observations of Contact Measurements Using a Resonance Based Touch Sensor

Marina Vidic, Salam M. Harb

School of Engineering, Coventry University, Coventry, UK
and

Stuart T. Smith

Precision Engineering Laboratory, UNC Charlotte, Charlotte, NC 28223

Introduction

Resonator based sensors have been proposed and used successfully in wide range of applications [1-3], which include measuring density and viscosity of gases or liquids, liquid level, sensing stress, pressure and mechanical forces. Although a considerable amount of work has been devoted towards the development of resonant touch sensors for scanned probe microscopy, commercial exploitation for precision applications has not been particularly successful. One of the reasons for this is due to the complexity of the interaction between a resonant sensor and typical surfaces in less than ideal environments. In this paper, a systematic study of the contact mechanisms between a resonator sensor and broad range of engineering surfaces are described. To do so, a resonator-based touch-sensor has been designed and manufactured and consists of a simple cantilever with actuators and sensors using PZT piezoelectric materials. Contact is detected by oscillating the sensor, which behaves like a second order system, near to its resonance frequency and monitoring the phase shift using phase locking techniques as the tip of the cantilever is approached towards a surface. The sensor could be employed for a broad range of applications including coordinate measuring machine probes, stylus transducers, friction measurement and surface mechanical probes (i.e. elastic modulus, density and energy dissipation). One particularly promising application is for monitoring changes in contact conditions or for the dynamic control of interface forces.

Sensor design and theoretical model

The resonator sensor used in the present investigation consists of a simple tip holder attached to one end of hexagonal BeCu rod (6x3.5 mm; $l = 20$ mm) and this is, in turn, mounted to a rigid base at the other. Mid way along the axis of the rod and symmetrically disposed are six piezoelectric elements. These form three pairs, each consisting of one actuator and one sensor on the opposing side. Consequently, drive and response signals can be monitored in terms of frequency and phase using standard phase locking circuits. In general any changes in force at the probe tip will effect the dynamics of the system. Such process can be modeled as a simple longitudinal oscillator of length, L , with the boundary conditions of added mass, stiffness and damping at the free end. Applying Rayleigh's approximate method to this system it is possible to derive a simple second order model with an undamped natural frequency given by

$$\omega_n = \sqrt{\frac{k_c + k_e}{M + m_e}} / 3 \quad (1)$$

where k_c , M , k_e , and m_e are the equivalent stiffness and mass of probe and the effective added stiffness and mass during contact, respectively.

In this discussion we shall separately consider the effects of added mass, stiffness and damping and assume that superposition applies. If any one of these dominates, it will also be shown that the effect can be systematically extracted from a series of three measurements. To visualize this Tables -I and II show the direction of the output from a phase locking amplifier for different types of loading (inertial, compliant or dissipative) on the sensor. In some cases, it is possible that two of the three influences will be insignificant. Under such conditions, it is desirable to identify the dominant effect. This can be deduced by systematically monitoring the changes in output with the probe set either side of, and at, the resonant frequency. Here, we shall consider two modes of detection commonly used for resonant sensors:

1. Monitoring frequency at constant phase shift between driver and sensor, hereafter called mode I
2. Monitoring phase at constant frequency, hereafter called mode II

In both of these modes, modeling is considerably simplified if small variations about a 90 degree phase shift are considered. In reality this is not easy to achieve and there is almost always a shift away from this ideal condition. Therefore it is necessary to consider the general effects of variations of the end condition on a second order system over a band of frequencies centered around the undamped resonance.

Table I: Direction of frequency shift for mode I operation of resonant sensor

Dominant factor	Phase angle in non contact state		
	$ \phi < 90 $	$ \phi = 90 $	$ \phi > 90 $
Mass	Increase	Increase	Increase
Stiffness	Decrease	Decrease	Decrease
Damping	Increase	0	Decrease

Table II: Direction of phase shift for mode II operation of resonant sensor

Dominant factor	Phase angle in non contact state		
	$ \phi < 90 $	$ \phi = 90 $	$ \phi > 90 $
Mass	Decrease	Decrease	Decrease
Stiffness	Increase	Increase	Increase
Damping	Decrease	0	Increase

Causes of parametric changes on resonant sensors

To simplify this discussion, only relatively large radius probes attached to a relatively stiff oscillator mount will be considered for two touch regimes, these being

1. Near surface phenomena
2. Mechanical contact

Near surface phenomena

As a resonating touch sensor approaches a surface a number of effects will be present which introduce a mechanical force. Such forces and their gradients will effect the frequency response characteristics of the probe. For non magnetic, clean materials in vacuum the only interactive forces will be those due to retarded and non retarded Van der Waals forces and electrostatic forces from surface charges. As evidenced by the rapid growth of the AFM industry, these effects can be significant for extremely compliant and sensitive probes. In our investigations such effects are considered negligible in relation to others that will be present during measurement of engineering surfaces in air. Under these conditions, it is likely that the dominant *near* surface forces will be those due to squeeze film damping and contaminant films on either the probe and/or the surface.

Surface films

The surfaces considered above are assumed to be perfectly clean and contaminant free. This is far from the case for many engineering surfaces and their typical environments. Commonly, surfaces will become covered by a condensed moisture film. As a consequence, there will be a surface tension force resulting in capillary action. It is speculated by Grigg *et al.* [4], that this leads to long range attractive forces of the order of tens of nanoNewtons and once the meniscus has been contacted will result in a force that can be calculated from the equation

$$F = \frac{4AR\gamma\cos(\theta)}{(1+x/d)} \quad (2)$$

where θ is the contact angle, d is the film thickness on both probe and specimen (assumed equal) and γ is the free surface energy (72 mJ m⁻² for water).

Squeeze film damping between a sphere and a flat

For relatively low velocities, as is the case for small oscillations at the frequencies used in the following experiments, the normal force, F , between a sphere and a flat surface separated a distance x is given by the equation [4],

$$F = \frac{3\pi\mu}{x^2 p_a r^3 (1 + r^2)^2} \frac{dx}{dt} \quad (3)$$

where, μ is the viscosity of the medium between sphere and flat, p_a is the ambient pressure and r is radius of sphere.

It should be noted that the force is both proportional to the velocity and inversely as the square of the separation between the two surfaces

Mechanical contact

In all of the following, it will be assumed that the oscillation amplitude is sufficiently small that 'bouncing' does not occur. In practice, this will only occur if there is a sufficient load at the contact. This load can be a combination of both the applied and self generated (i.e meniscus and adhesion forces) components.

Effective stiffness (k_e)

From Hertzian analysis, it is relatively straightforward to show that the local stiffness for small oscillations of an elastic contact is given by the familiar equation [6],

$$k_e = (6E^2 PR)^{1/3} \quad (4)$$

where E , P and R are Young's modulus, contact force and probe radius, respectively.

Effective mass (m_e)

Upon physical contact with a material the region of the interface must be moving at the same frequency as the motion of the probe tip and will thus add an inertial load. Clearly, this will extend into the specimen with diminishing amplitude. To assess the effective added mass, it will be assumed that the specimen is smooth and flat and the probe is also smooth, has an effective radius, R . For velocities of the probe much less than those of stress waves in the contact materials, it is reasonable to assume that the strain field is the same as that derived from Hertz analysis. For a probe having an elastic modulus of considerably higher value than that of the specimen, the effective mass for small oscillations can be approximated as

$$m_e \approx \frac{\pi^2 PR}{2E} (1 - \nu^2) \rho \quad (5)$$

where ν and ρ are Poisson's ratio and density of the material, respectively.

This equation indicates that the added mass becomes significant for materials with high density and low modulus of elasticity e.g. natural rubber. Fig. 1, shows the theoretical phase shift in the resonant probe for a selection of different materials under various loads, which was calculated using equations (1), (4) and (5) and estimated values for damping and coupling loss factor (ζ).

Effective damping (h_e)

Most compliant materials such as rubber and plastics can be characterized as exhibiting dissipation characteristic of hysteretic damping. Mathematically, this can be modeled as complex stiffness (often referred to as structural damping). A second order system with hysteresis damping, h_e , subject to sinusoidal input can be characterized by the equation:

$$m\ddot{x} + k(1 + ih/k)x = F_0 e^{i\omega t} \quad (6)$$

Assuming a solution of the form $Xe^{i\omega t}$, and also that the hysteresis is predominantly due to the cyclic adhesion between the two surfaces, h_e can be solved to give:

$$h_e = 1.23\gamma_{12} \left[\frac{R^4 P}{E\Delta x^6} \right]^{2.9} \quad (7)$$

where, γ_{12} and Δx are the interfacial energy characterized by the Dupré equation and amplitude displacement of the probe at the interface respectively.

Experimental results

To assess the interaction between a probe tip and typical engineering surface the resonator touch probe was mounted onto a simple linear flexure stage for fine positioning, Fig.2. This was then positioned horizontally, in line with a stationary specimen. The contact force between the probe tip and a specimen was monitored using two strain gauges in a half bridge arrangement, mounted onto the specimen. During these experiments, the phase shift in mode I operation (or frequency shift for mode II operation) and applied forces were recorded as the probe approached and contacted the surface.

Contact characteristics of thick oil film

Experiments were conducted to assess the influence of a thick oil film with the reference frequency or phase set to each side of the free resonance. Fig.3 shows a typical plot of the probe's phase shift from the free value (non contact) set to -81 degrees, Fig. 3a., and -100 degrees, Fig. 3b, as it was contacted with a thick layer of silicone oil having a relatively high viscosity of approximately 10 Pa s. A number of distinct regions could be seen on the graphs. The first region corresponds to approach of the probe tip towards the oil surface prior to physical contact where there is an increase in phase in both figures which from table II indicates a predominant inertial force. The immediate contact with the oil results in the reduction in phase in Fig. 3b, which is not present in Fig. 3a. Referring again to Table II, it is possible to explain this as an effective damping which would add to the inertial effect with opposite sign either side of the resonant frequency. Upon further displacement of the probe into the oil the phase shift increases and it could be observed that in Fig. 3a, it has opposite curvature to the one in, Fig. 3b, but at a faster rate for phase less than -90 degrees. Referring to the Table II, this suggests that if the free value is below resonance, the two predominant effects, mass and damping, are summed, while the value above the free resonance results in their difference. As the probe approaches the surface the phase increases to a value above |90| degrees at which point an additional damping results in a reduction of the phase. When mechanical contact takes place, phase rapidly reduces as a result of a large increase in effective stiffness. Retracting the probe away from the surface, the phase shift mirrors the original approach curve although there is a permanent offset of +2 degrees. Cleaning of the probe after separation from the film returned the phase to the original free value of -81 degrees, or -100 degrees.

Contact of solid surfaces

A number of experiments were conducted on different type of materials: PTFE, aluminum, copper, steel, glass and two types of rubber. Fig. 4 shows a typical plot of a probe's phase shift in mode II operation when contacted with these solid surfaces. Prior to each experiment the surface were thoroughly cleaned using acetone. In Fig. 4a, and 4b the free value phase shift was set at -85 degrees and -100 degrees, respectively. For PTFE, the phase shift starting at less than |90| degrees was decreased of less than one degree, whereas for glass it was decreased of more than 25 degrees, for the same contact force. When this compared with similar test but for phase shift starting from above |90|, materials like PTFE resulted in more or less similar decrease in phase of less than one degree whereas for glass it was decreased about 20 degrees. Close comparison with other materials tested indicates that, on average, the difference in the phase shift for the same material is about 5 degrees when the phase is below and above |90| degree. Again, referring to Table I, this suggests that this is the contribution of the added damping in the system although these large phase shifts may include other non symmetric factors. As predicted, all the materials tested in Fig. 4 show a reduction in phase, but at different rates, which indicates that added stiffness is the predominant effect. In contrast, similar tests conducted with rubber show an increase in phase which indicates the relatively high density and low elastic modulus results in added mass being the predominant effect. Fig.5.

Discussion and conclusion

It has been shown that the dynamics of a resonant probe sensor as it approaches typical engineering surfaces is subject to a complex variety of forces leading to the observation of effects characteristic of inertial, stiffness and damping forces. These characteristics will superpose making the separation of individual terms rather complex. However, it has been shown that observation of dynamic characteristics either side of the resonant frequency may provide a systematic method to identify the predominant effect. Preliminary theoretical and experimental investigations on both contact of thick oil films and solid surfaces provide encouraging results.

It is also speculated that further analysis of oscillatory contact with solid surfaces and comparison with corresponding hardness scales could lead to potential applications in non-destructive hardness testing. That the shift in the resonant frequency (or phase shift) depends on near surface properties such as the elastic modulus, hardness, density and internal damping makes the quantitative measurements of a number of physical properties an exciting possibility.

Acknowledgments

The authors are grateful to Coventry University and the Nuffield Foundation for supporting this research.

References:

- [1] Langdon R.M. Resonator sensors-a review. *J. Phys. E: Sci.Inst.*, **18** (1985), 103-115.
- [2] Morten, B., De Cicco, G., Prudenenziati, M. Resonant pressure sensor based on piezoelectric properties of ferroelectric thick films, *Sensors and Actuators A*, **31** (1992), 153-158.
- [3] Omata, S. and Terunuma, Y. Development of a new type tactile sensor for detecting hardness and/or softness of an object like the human hand, *IEEE*, 1991, 868-971.
- [4] Gross W. A., 1962, *Gas Film Lubrication*, John Wiley and Sons, NY, p325
- [5] Grigg D.A., Russell P.E. Griffith J.E. Tip-sample forces in scanning probe microscopy in air and vacuum *J.Vac.Sci. Technol.*, **A10**(4) 680-683.
- [6] Smith, S.T. Chetwynd D.G. *Foundations of Ultraprecision Mechanism Design*, Gordon and Breach, 1992, NY.

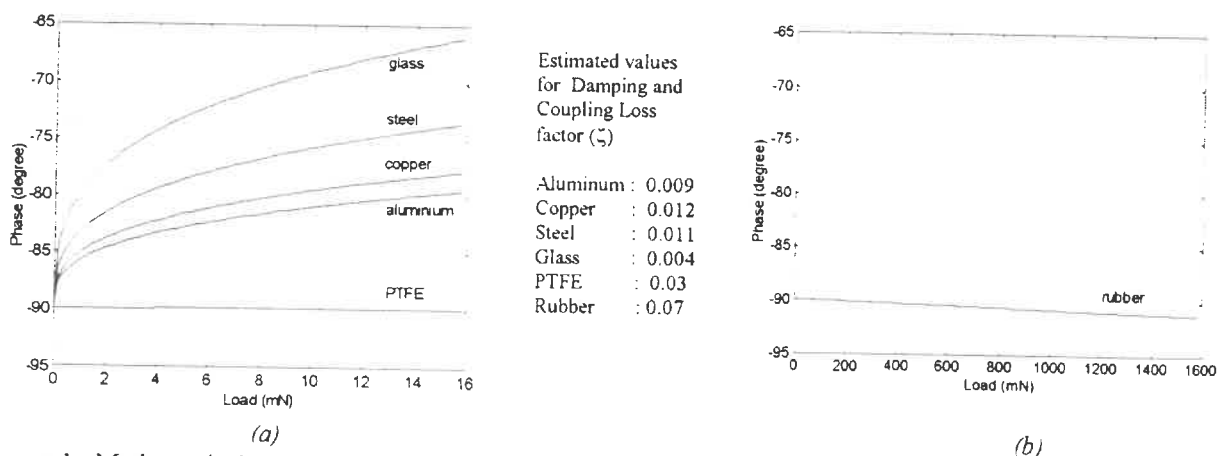


Figure 1. Mathematical results representing the phase shift in a resonant probe for a selection of different materials as a function of applied load and for given probe geometry.

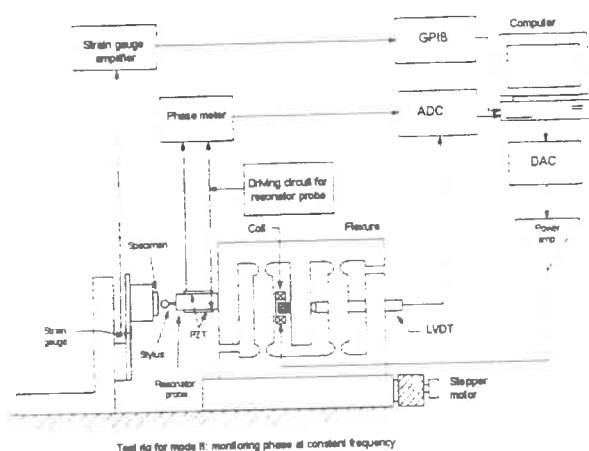


Figure 2. Schematic diagram of the test rig used to monitor the contact between resonator probe and solid surfaces

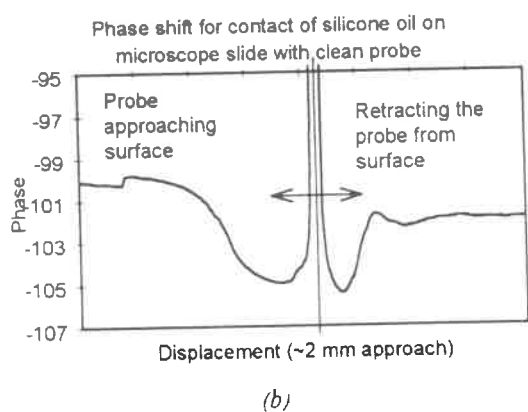
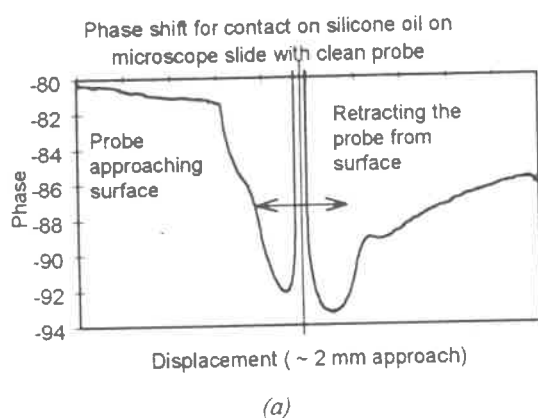


Figure 3. Plot showing contact characteristics of thick oil film for phase shift of a) -81 degrees and b) -100 degrees

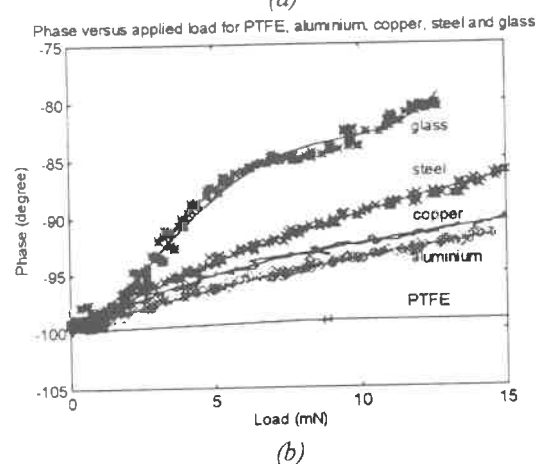
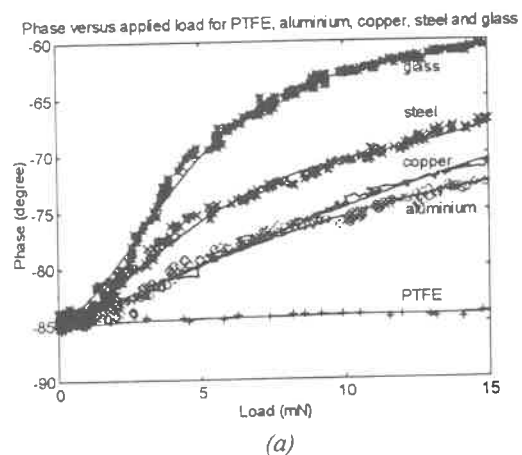


Figure 4. Plot showing phase shift when probe contacts solid surface for free shift of a) -85 degrees and, b) -100 degrees.

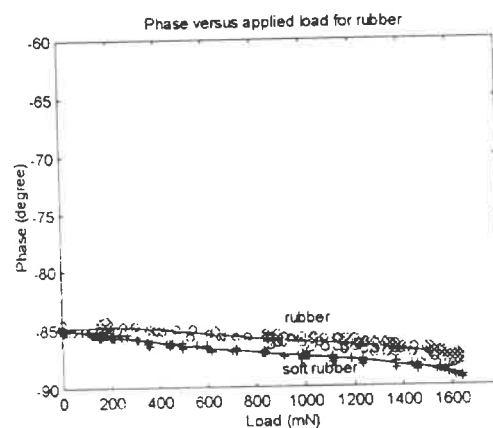


Figure 5. Plot showing phase shift when probe contacts rubber like materials

Novel Linear Encoder Incorporating Incremental Optical and Absolute Capacitive Encoders

Shawn A. Hogan and Karl G. Masreliez

Micro Encoder Inc.

Kirkland, WA 98034

1. Introduction

In this paper we describe a novel new absolute encoder, referred to as a "combo" encoder, developed by Micro Encoder Inc. for Mitutoyo. The combo encoder incorporates the benefits of incremental optical and absolute capacitive encoders into a compact, robust encoder for the machine tool and motion control industries.

Continuing advances in the machine tool and motion control industries place greater demands on the position feedback encoders. Systems require higher resolution, for increased accuracy, with increased system velocity. The encoders are often located in hostile environments, where they are subject to mechanical shock, electrical noise and unstable power supplies. To support future speed requirements and improve encoder robustness, it is beneficial to use absolute encoders for the position feedback.

Traditional methods of incorporating absolute encoder designs using optical techniques require multiple tracks. The space required to implement such a design and the detector and signal processing circuitry make this option both expensive and complex. Recent advancements in absolute position detection use capacitive detectors, e.g., digital calipers. These detectors implement the absolute position detection in both a small area and with minimal signal processing. However, it is difficult to achieve high resolution in a capacitive detector. By combining the two detection methods, we obtain a system with both high resolution and absolute position determination in a compact form.

2. Scale Mechanics

The orientation of the optical and capacitive detectors is shown in Figure 1.

The capacitive detector is split into two sections, with the optical detector located at its centroid. Abbé and thermal expansion errors are minimized by locating the detectors in the manner.

The optical (OPT) scale grating is integrated into the capacitive fine (FINE) scale electrodes, as shown in Figure 2. The scale elements are made of thin film chrome on glass, and are fabricated by photolithography methods.

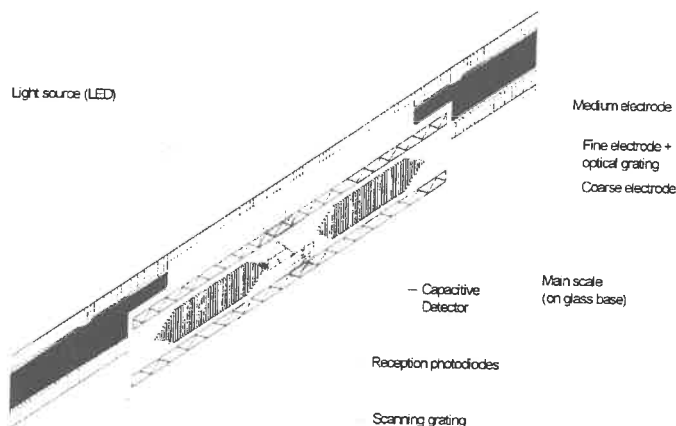


Figure 1: Orientation of optical and capacitive detectors¹

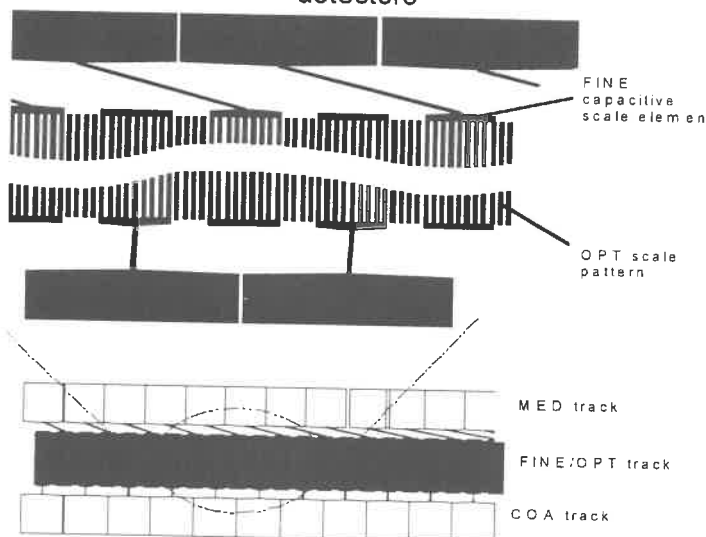


Figure 2: Detail of scale pattern²

The detector assembly, shown in Figure 3, is attached to the scale mounting block using three flexures and an intermediate support. The flexures provide a compliant link between the detector assembly and the scale mounting block, so that mechanical disturbances incident upon the scale block are not transmitted to the detector assembly.

The detector assembly is kinematically supported against the surface of the glass scale by five ball bearing wheels and is preloaded against the scale by two preloaded flexure springs with integral ball bearing wheels. The ball bearing wheels and preload flexures constrain five degrees-of-freedom for the detector assembly, only permitting linear motion along the scale axis.

3. Principle of Absolute Position Detection

The encoder has four distinct measurement wavelengths or ranges. The absolute capacitive detector provides the three upper ranges: FINE, medium (MED) and coarse (COA). The optical detector provides the finest or OPT range. The absolute position is determined from measurements and interpolation of the various ranges. The COA range is used to determine which of many similar MED ranges the encoder is within. The MED and FINE range determine the particular incremental OPT range. Finally, interpolation of the OPT signal determines the absolute position within $1\mu\text{m}$. Figure 4 illustrates the measurement process. The relationship

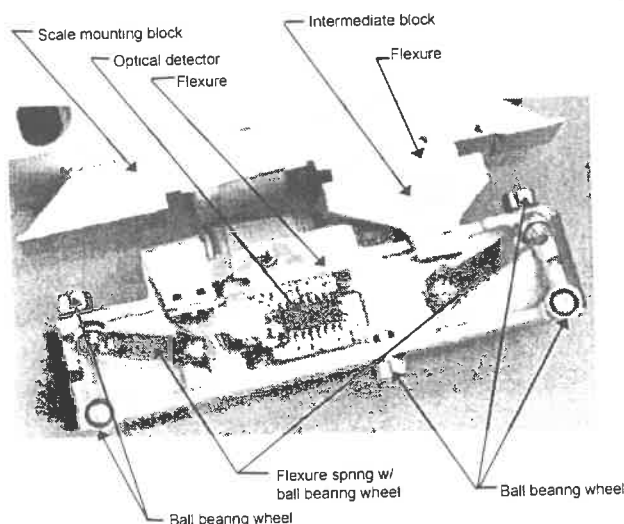


Figure 3: Encoder detector assembly

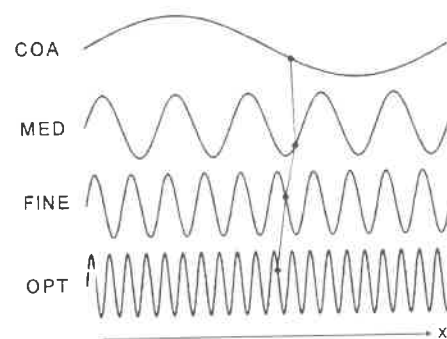


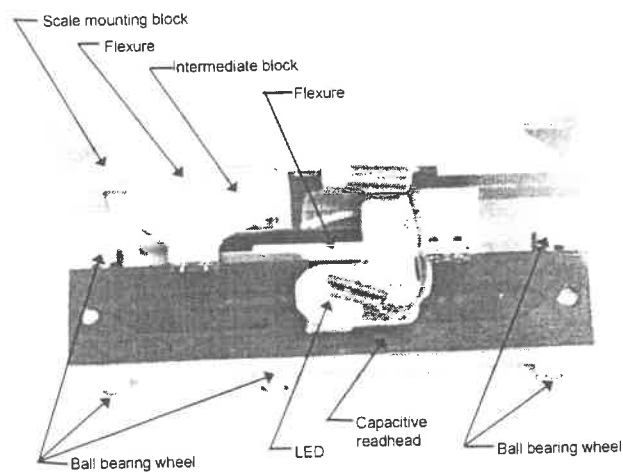
Figure 4: Absolute measurement process³

between the four ranges in the combo encoder are shown in Table 1.

Range	Symbol	Relationship
OPT	λ_o	0.040 mm
FINE	λ_f	1.040 mm
MED	λ_m	$40 \cdot \lambda_f = 41.600 \text{ mm}$
COA	λ_c	$40 \cdot \lambda_m = 1600 \cdot \lambda_f = 1664.000 \text{ mm}$

3.1 Optical Incremental Encoder

The optical encoder is a transmission-type optical detector, with the scanning grating located in front of the reception photodiodes⁴. The scale consists of a series of clear and opaque bars formed on a glass base (see Figure 2). The period of the scale patterns is $40\mu\text{m}$, with 50% duty cycle. A light source is incident upon the patterned scale from one side of the scale. The transmitted light forms a periodic array of light and dark regions upon the scanning grating. The scanning grating consists of four windows, patterned similar to the scale, with the gratings offset by $10\mu\text{m}$ relative to each other.



Four photodiodes are arranged in an array behind the scanning grating.

The result is that the current through the photodiodes will vary depending upon the position of the scale relative to the readhead. This provides output signals proportional to $\pm\cos(kx)$ and $\pm\sin(kx)$, where k is $2\pi/40\mu\text{m}$, the wavenumber of the optical detector. Using these four signals, the OPT position is determined to $1\mu\text{m}$ resolution by 40X interpolation of the analog photodiode signals.

3.2 Capacitive Absolute Encoder^{5,6}

The capacitive encoder is comprised of two parts, a scale and a readhead, as shown in Figure 5. The scale is characterized by three interdigitated electrode tracks. Alternating electrodes in the center track are connected to the outer tracks. The scale has no external electrical connections.

The capacitive readhead features three electrode arrays and all electrical connections. In operation, the readhead faces the scale, and the array of electrodes form a series of parallel plate capacitors between the two parts.

An electrical signal transmitted from the center readhead electrode array, or FINE transmitter (T_n), is capacitively coupled to the center scale track, or FINE receiver (SR). The center scale track is connected to the tracks on either side and the electrical signal is, therefore, present on the outside tracks as well, the MED and COA transfer electrodes (ST_m and ST_c). The signals on ST_m

and ST_c are capacitively coupled back to the outer readhead electrodes, or MED and COA receivers (M_1 , M_2 , C_1 and C_2). The magnitude of the capacitive coupling between the elements changes as the scale moves relative to the readhead. The voltage on the receivers (V_{M1} , V_{M2} , V_{C1} and V_{C2}) is, therefore, a function of the relative position of the scale and readhead.

3.2.1 Fine Measurement Process

In the fine measurement process, the two halves of the MED receiver, M_1 and M_2 , and the two halves of the COA receiver, C_1 and C_2 , are summed in the electronics, and we may treat the two receivers as a solid electrode of uniform width.

As the scale moves relative to the readhead, the capacitance function between a particular transmitter, T_n , and alternating SR electrodes (either those connected to ST_m or ST_c) varies as a function of position.

Using the MED transmission path as our example, we can express the capacitance function as

$$C_n(x) = C_0 \sin\left(\frac{2\pi}{\lambda_f} x - \frac{2\pi}{N} n\right) + \varepsilon$$

where x is the relative displacement

ε is some unknown offset

and N is the number of transmitter phases, typically eight.

In addition, the signal applied to the transmitters is modulated, depending upon the transmitters' relative location on the readhead. Therefore, the applied signal on the n^{th} transmitter is

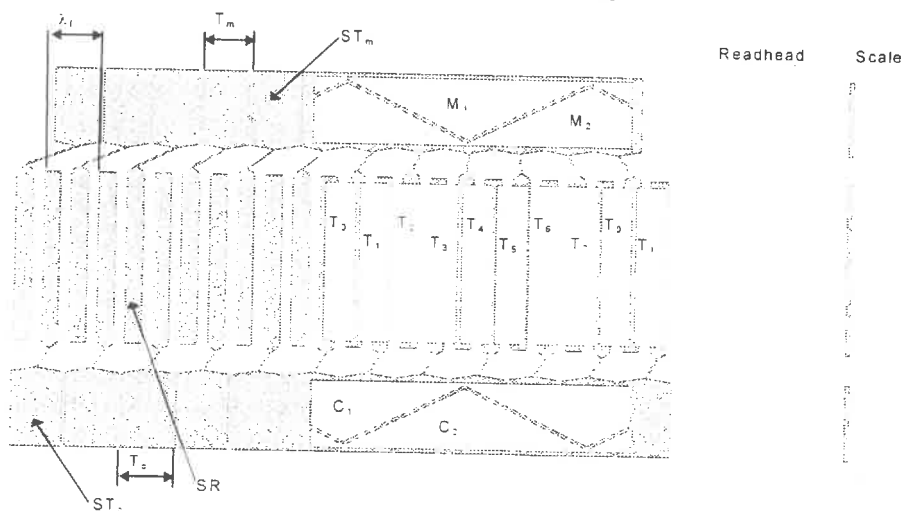


Figure 5: Capacitive encoder electrode arrangement

$$V_n(t) = V_0 \sin\left(\omega t - \frac{2\pi}{N}n\right)$$

The signal on the receiver, due to the n^{th} transmitter, will be proportional to the product of the input signal and the capacitance function. The total signal on the receiver will be proportional to the sum of the individual transmitter signals

$$\begin{aligned} V_{EM} &= V_{M1} + V_{M2} = \sum_{n=0}^{N-1} A \cdot C_n(x) \cdot V_n(t) \\ &= A \cdot C_0 \cdot V_0 \sin\left(\frac{2\pi}{\lambda_f}x - \omega t\right) \end{aligned}$$

where A is a proportional constant. We see that the signal on the MED receiver is indeed periodic, with displacement in x .

For the COA receiver, however, the SR electrodes connected to the ST_c track are 180° out-of-phase with the SR electrodes connected to the ST_m track. Therefore, by subtracting the two values, we find the fine position from

$$\begin{aligned} V_F &= V_{EM} - V_{EC} = (V_{M1} + V_{M2}) - (V_{C1} + V_{C2}) \\ &= 2A \cdot C_0 \cdot V_0 \sin\left(\frac{2\pi}{\lambda_f}x - \omega t\right) \end{aligned}$$

3.2.2 Medium and Coarse Measurement Process

In the MED and COA measurement process, the two halves of the MED receiver, M_1 and M_2 , and the two halves of the COA receiver, C_1 and C_2 , are subtracted in the electronics. The MED and COA measurements take advantage of the relationship between the SR and the ST_m or ST_c electrode arrays. Using the MED measurement process as our example, we make the pitch of the ST_m electrodes, T_m , related to λ_f by a ratio, usually less than one, i.e.,

$$T_m = \frac{m}{m+d} \lambda_f = \frac{40}{40+5} \lambda_f$$

where m is the MED ratio, 40 for the combo and d is the scaling ratio.

The addition of d unconnected electrodes into the ST_m array for every m ST_m electrodes connected to the SR array results in a periodic relationship between the SR and ST_m electrodes. For every $(m+d)$ ST_m electrodes, the relative offset between a common point in the SR array and the ST_m array is the same, giving a MED range of

$$(m+d)T_m = (m+d) \frac{m}{m+d} \lambda_f = 40\lambda_f$$

A similar argument is made for the COA measurement process, with the substitution of $c=1600$ for $m=40$.

4. Conclusion

We have discussed the design and operation of a new high-resolution, absolute linear encoder, characterized by the use of two forms of position detection, a high-resolution incremental optical detector and an absolute capacitive detector.

The incremental optical detector is a transmission-type detector, with the scanning grating located in front of the photodiodes. The output from the optical detector is in analog quadrature, interpolation of which provides $1 \mu\text{m}$ system resolution.

The absolute capacitive detector is of the type developed for absolute digital calipers. It provides position information at three characteristic wavelengths or ranges. Measurement and interpolation of the three capacitance ranges provides information as to which of many optical ranges the encoder is within, providing the absolute encoder position.

The detector is kinematically mounted against the scale using five ball bearing rollers. The detector assembly is attached to the scale mounting block by means of a series of flexures, which insulate the detector assembly from external mechanical disturbances.

¹ Kiriya, T., Nakayama, K., and Yamaguchi, Y., "Development of Absolute Linear Scales AT41/AT35", *Mitutoyo Technical Bulletin*, No. 14, 1996.

² *Ibid.*

³ *Ibid.*

⁴ Walcher, I.H., *Position Sensing Angle and Distance Measurement for Engineers*, Trans. Kerr, D. and Shields, M.J., Oxford: Butterworth Heinemann, 1994, pp. 86-90.

⁵ Andermo, I.A., Shimomura, T., Kiriya, T., and Yamaguchi, Y., "An Absolute Linear Encoder Utilizing a Combination of Capacitive and Optical Encoder Technology", *Japan Soc. Prec. Eng.*, Vol. 29 (1), March 1995

⁶ T., Kiriya, Shimomura, Andermo, I.A., and Yamaguchi, Y., "Development of Absolute Linear Encoder", *Japan Soc. Prec. Eng.*, Vol. 61 (10), 1995

A New Design for a High Speed Spindle

M. Weck, S. Fischer	Fraunhofer-Institut für Produktionstechnologie IPT, Aachen, Germany
P. Holster	Philips Research Labs, Eindhoven, The Netherlands
K. Carlisle	Cranfield Precision Ltd, Cranfield, England
Y. Chen	Carl Zeiss, Oberkochen, Germany

Introduction

Precision grinding and micromachining both imposes high demands on the machine behaviour, since the achievable workpiece accuracy is determined not only by the technological parameters but also by the characteristics of the applied machine components.

Ultraprecision surface quality and the mechanical fabrication of structures in the micron range can only be achieved by using machine tools which have appropriate spindles. Structures cannot be manufactured using spindle types of which the radial error motion is greater than the level of contour accuracy or surface roughness required.

In addition, the spindle speed is an important value. Not only a certain cutting speed is needed from the technological point of view, but also the machining time required for microstructuring surfaces is reduced by deploying a high frequency spindle, thereby increasing the economic efficiency of the technique.

Hence, the main purpose of the project was to develop a high speed spindle with properties concerning accuracy, speed and stiffness beyond commercially available ones.

Spindle Design

The airbearing spindle is built as a hybrid design. This bearing type combines spiral groove technology, known from aerodynamic bearings, with an externally pressurized bearing. The spindle is driven by an integrated motor with the rotor fixed directly on the spindle shaft, enabling a maximum speed of 100.000 rpm.

High friction losses in the bearings and the power loss of the motor require a water cooling system, integrated into the spindle. As shown in further tests, the spindle is thermally insensitive, both in axial and radial direction, within a tolerance range of $0.5 \mu\text{m}/^\circ\text{C}$.

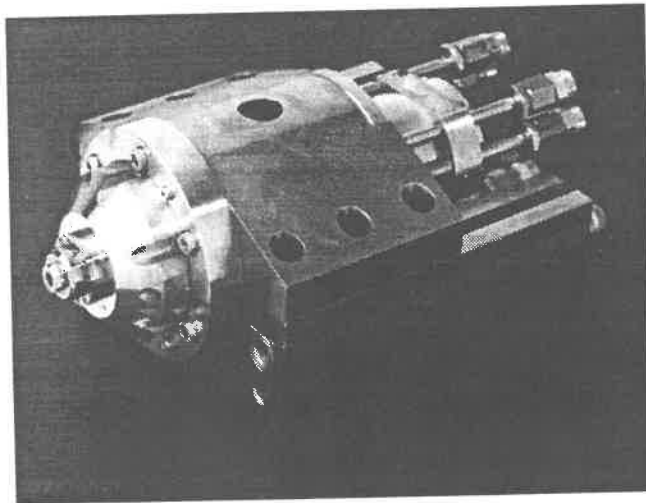
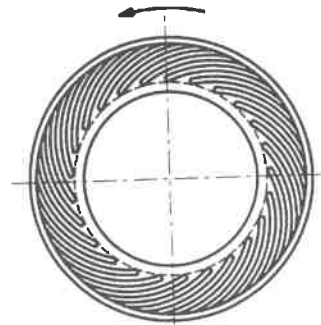
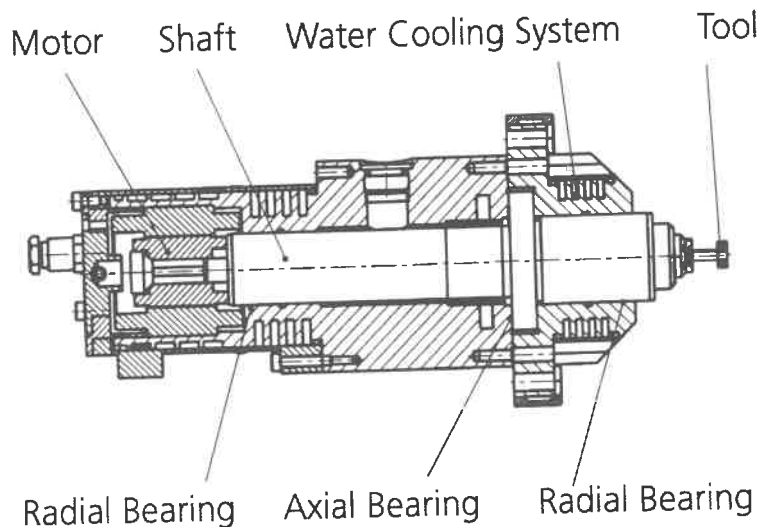


Fig. 1: High Speed Spindle Design

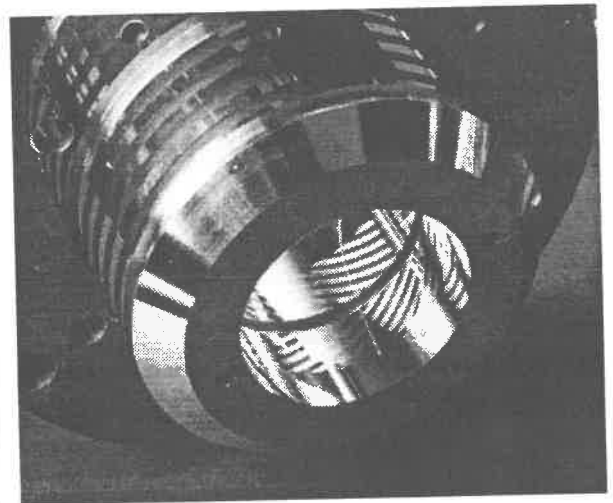


Fig. 2: Spiral Groove Structure of the Axial Bearing (depth: 5 μm)

Fig. 3: Prototype of the High Speed Spindle

Fig. 3: Radial Bearing (Front) including Spiral Grooves and Cooling Pipe

To predict the spindle deformation due to centrifugal forces, thermal expansion and aerostatic pressure, a corrected gap geometry was calculated and manufactured. Conical parts of 3 μm depth were machined into the radial bearings to achieve a constant bearing gap width of 5 μm at maximum speed.

Spindle Measurements

Measurements of the dynamic spindle behaviour were taken at stand-still and at rotating speed. From the following compliance frequency response curve G_{xx} , taken at stand still, it can be clearly seen, that the compliance never reaches the static value at frequencies below 1900 Hz., which implies a noncritical dynamic behaviour

of the spindle in the whole speed range. In addition, sharp peaks in the compliance curve were not found. This leads to the conclusion that the resonance positions are very well damped.

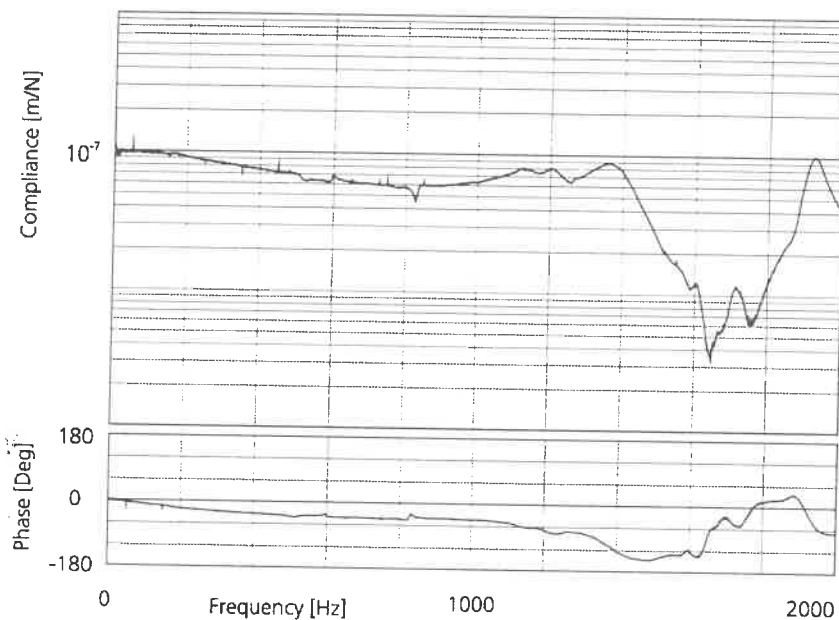


Fig. 5: Compliance Frequency Response Curve Gxx

In addition, the bearing stiffness was measured at rotating speed by machining an aluminium master on the spindle at different speeds. During the cutting process, the tool was excited with a piezo actuator, implying a dynamic force (white noise signal) into the spindle. Simultaneously, the displacement of the master was measured. The analysis of the force and displacement signals was executed with a Fourier-analyzer. At maximum spindle speed, a static stiffness of 30 N/ μ m is obtained at the cutting point, rising from about 10 N/ μ m at stand still.

It emerged, that the radial and axial total error motion of the spindle is below 200 nm during the whole speed range. At spindle speeds up to 10.000 rpm, a total error motion of 40 nm is observed

Machining Results

Grinding tests using the hybrid spindle were performed under workshop conditions. The achieved results with test pieces made of steel and glass demonstrated the excellent properties of the spindle. Surface qualities below 2 nm Ra were achieved. In addition, diamond milling experiments were carried out. It emerged, in the very first tests conducted on an 3-axis ultraprecision machine, that microstructures can be produced with structure sizes far below previously known values. The thin ribs, shown in figure 6, were machined into brass using a diamond milling tool with a

diameter of 300 μm . Minimum rib sizes of 1.5 μm width x 200 μm height were obtained.

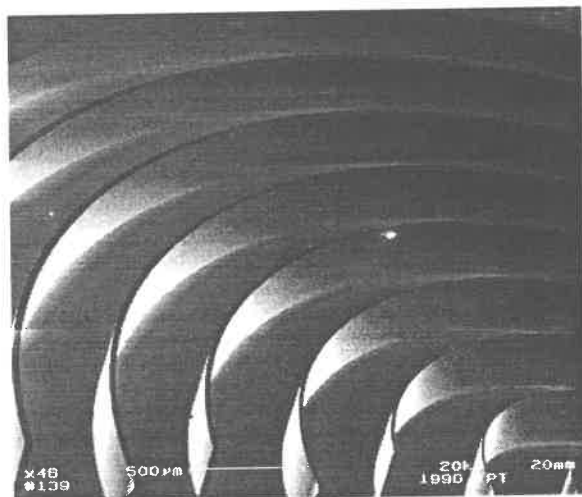


Fig.6: Micro-Ribs, Milled into Brass

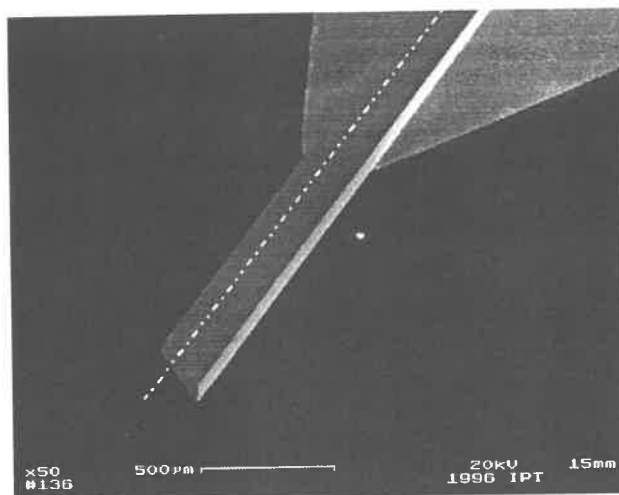


Fig. 7: Diamond Milling Tool, 300 μm dia.

Conclusion

A high speed spindle was developed, combining an externally pressurized air bearing with a spiral groove design, known from aerodynamic bearings. The spindle properties concerning accuracy, speed and stiffness are far beyond commercially available ones.

The excellent qualities of the spindle were demonstrated by the obtained machining results. Surface roughnesses below 2 nm Ra were achieved by grinding. Diamond milling experiences showed, that micro-ribs with minimal sizes of 1.5 μm x 200 μm can be machined into brass.

References

1. Holster P., Jacobs J., Roblee J. „The Measurement and Finite Element Analysis of the Dynamic Stiffness of Non-Uniform Clearance Gas Thrust Bearings“, ASME Journal of Tribology, 1990
2. Weck M., Hartel R., Bispink T., Luderich J. „Konstruktion, Berechnung, Messtechnische Untersuchung und Einsatz einer luftgelagerten Arbeitsspindel für eine Ultrapräzisionsdrehmaschine“ Review at the Airbearing Seminar TA Esslingen, Germany, 1989
3. Roblee, J. „Precision Temperature Control for Optics Manufacturing“ SPIE Cannes, France 1985
4. J. Franse, de Jong G. „Generation of Surface roughness in Precision grinding of Small Molds, Proc. SPIE, Vol 803, pp. 43-52

PRECISION GRINDING OF BRITTLE MATERIALS: THE ASPE 1996 TOPICAL MEETING - SUMMARY AND REVIEW

*R. O. Scattergood
Precision Engineering Center
Raleigh, NC 27695 - 7918*

This presentation will differ from the usual technical format in that it will be primarily a summary and review of the highlights from the ASPE Spring Topical Meeting held June 3-6, 1996 at St. John's College, Annapolis, MD. In that context, this will be a short (as opposed to extended) abstract. The theme of the meeting was Precision Grinding of Brittle Materials, a topic endemic to the precision engineering and manufacturing of a wide range of components. The meeting was organized with the aim of bringing together a representative group of participants from industry, national laboratories and academia. The sessions included issues related to machine design, grinding wheels, coolants, materials and the mechanics of the grinding process as manifest in chip geometry and material removal mechanisms. The rather sharply focused content of the meeting evoked numerous lively and sometimes controversial discussions on what we know, what we don't know and what we should know about the grinding of brittle materials. A summary session led by K. L. Bladel helped put much of the technical detail into perspective. The meeting was successful in producing a better picture of where we stand in terms of a fundamental understanding of the grinding process as well as our ability to transfer this understanding into more deterministic manufacturing technology. Perhaps most interesting was the realization of the still fuzzy nature of our current state of knowledge in many important areas, even after many years of research by many investigators. That itself can be judged as an advance in that at we are beginning to know better what we don't know. The rest of the details will be filled in at the presentation.

Diamond Machining of the 3m Reflector of the KOSMA Submillimeter Telescope by a Single-point Fly-cutting Process

E. Brinksmeier, W. Preuss

LFM Labor für Mikrozerspanung, University of Bremen, Germany

Abstract

We have developed a point contact fly-cutting process for diamond machining of asymmetric mirror surfaces. The method was applied for finishing the 18 segments of the parabolic reflector of the 3m KOSMA submillimeter telescope operated by the University of Cologne on the Gornergrat (Zermatt) in the Swiss Alps.

1 Introduction

Since its beginnings in the 1930s, radio astronomy has developed into a most powerful scientific tool which has revolutionized our picture of the universe. The range of the radiowave window of the atmosphere, limited by ionospheric reflexion at the long wavelength end and by aerosol and molecular absorption at the lower end, extends over nearly five orders of magnitude (cf. **fig. 1**).

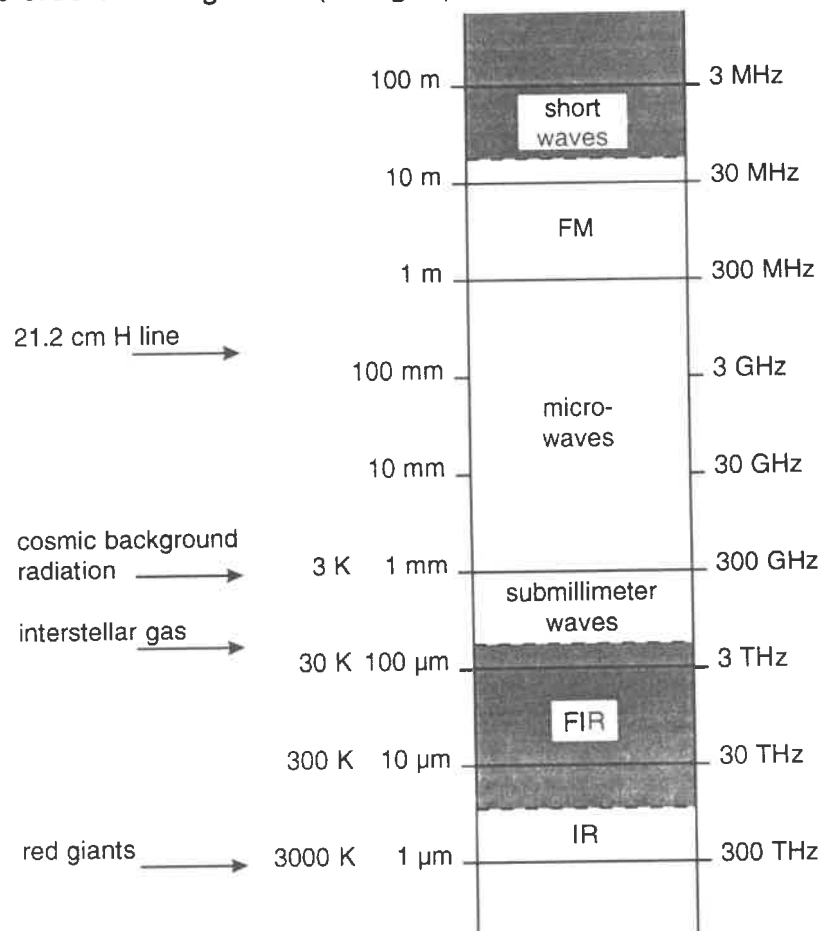


Fig. 1: Radiowave window of the atmosphere.

While most non-thermal radio sources, e. g. pulsars, quasars, cosmic masers and supernova remnants, are observed in the FM and microwave bands, the dynamics of interstellar gas clouds eventually evolving into new born stars is studied in the submillimeter band. In recent years, along with the development of liquid helium cooled detectors, attempts have been made to push the frontiers of submillimeter astronomy into the high frequency range beyond 500 GHz. Today's most advanced submillimeter telescope (SMT) is the 10m Heinrich-Hertz telescope erected on Mount Graham, Arizona, jointly operated by the Steward Observatory of the University of Arizona and the Max-Planck-Institute for Radioastronomy, Germany. Racing for higher frequencies, the 3m KOSMA SMT on the Gornergrat in Switzerland is now being upgraded by the University of Cologne for observations at frequencies up to 1 THz.

The shortest wavelength at which a SMT can be operated with good resolution is determined by the total error budget of the telescope which is defined as the root-mean-squares sum of contributions from the deflection of the backup structure, thermal drifts, alignment errors and - last not least - figure errors of the parabolic primary reflector. The goal for the KOSMA SMT is a 10 μm total error budget imposing a new challenge for machining the primary reflector with a focal length of 1.3 m to within a 5 μm rms figure tolerance. The reflector is composed of 18 light-weight aluminum segments arranged in two rings (cf. **fig. 2**). Due to their large off-axis distances and deep asphericity, the segments cannot be turned off-axis, even on the largest existing precision lathes, nor can they be machined on-axis by means of a fast tool servo.

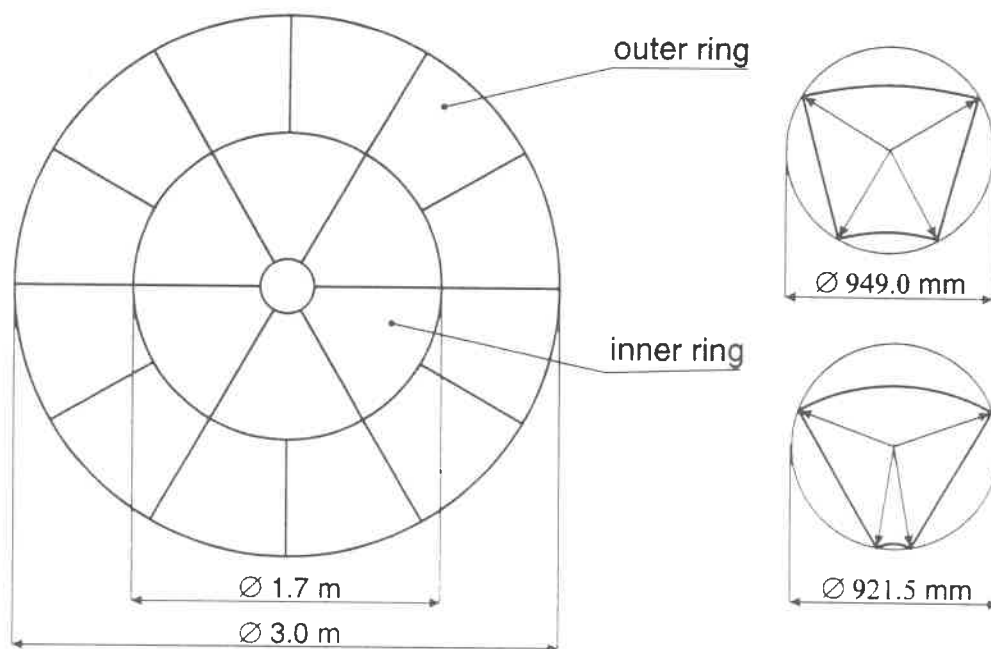


Fig. 2: Arrangement of the 18 panels of the KOSMA primary reflector.

2 The single-point fly-cutting (SPFC) concept

Machining of asymmetric surfaces involves a minimum of three programmable machine axes. Takeuchi et. al. [1] have reported the generation of elliptic paraboloid and toroid surfaces by means of a 3-axes ultra-precision ball-end milling process. The method, however, becomes uneconomic if large surfaces have to be worked. Falter and Dow [2] and Weck and Pyra [3] have been the pioneers of the fast tool servo approach which yields high quality asymmetric and even discontinuous surfaces whose departure from a rotationally symmetric surface, however, has so far been limited to a couple of microns.

We have replaced the piezo-driven tool of a fast tool servo set-up by a diamond fly-cutter which is moved back and forth by means of the z-slide of the machine according to the angular position of the slowly rotating workpiece (cf. **fig. 3**). Chips are removed at high cutting speed due to the fast rotation of the fly-cutter. Large depths of cut can be realized, since the interaction time between tool and workpiece is extremely short and the torque on the fly-cutter is negligible. We have termed this process single-point fly-cutting (SPFC) in contrast to ordinary fly-cutting which is a line contact process. As a matter of fact, SPFC is by no means a novel technique but has widely been used for fly-cutting of grooves, gratings und prisms. However, SPFC is applied here for the first time in a 3-axes operation for generating asymmetric surfaces.

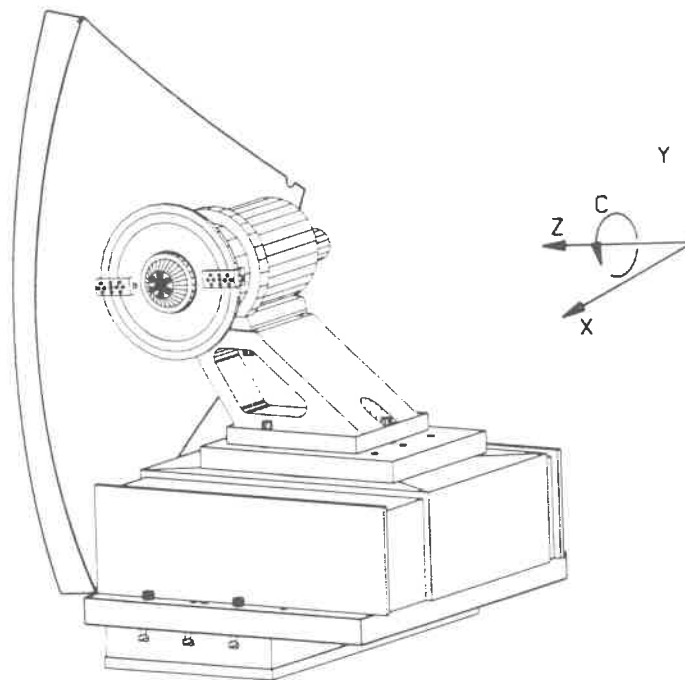


Fig. 3: Set-up for single-point fly-cutting of reflector panels.

Due to the point contact nature of SPFC, the machined surfaces exhibit tiny pits or “scallop” whose width w_c along the cutting direction depends on the frequency ν of the fly-cutter and the radial velocity v of the rotating workpiece and whose width w_f in feed direction equals the feed f per revolution of the workpiece:

$$w_c = v / \nu , \quad w_f = f .$$

The depths of the pits in the cutting and feed direction, h_c and h_f resp., are given by the widths w_c and w_f and by the fly-cutter radius r_c and diamond radius r_d :

$$h_c = w_c^2 / 8r_c , \quad h_f = w_f^2 / 8r_d .$$

r_c and r_d must be smaller than the radii of curvature of the surface to be machined. The size of the pits can be reduced by lowering the feed and/or the radial velocity of the workpiece which - on the other hand - increases machining time. The periodic pit pattern may lead to diffraction of infrared radiation and must carefully be optimized in order to meet optical requirements. However, for submillimeter applications even a coarse pit pattern is tolerable.

The process can be run in the up-milling or down-milling mode. Tool life in SPFC is surprisingly long, particularly in the up-milling mode, as compared with SPDT. This might be explained by the fact that the average cutting time per unit volume of cut material is much smaller in SPFC than in SPDT.

3 Cutter compensation

Since the effective shape of a rotating diamond tool is a toroid, cutter compensation must take into account the shift of contact point of a toroidal tool machining an asymmetric part. If the surface is given by a differentiable function

$$z = f(x, y),$$

the gradient $\nabla f(x, y)$ is defined at any point of the surface and the true contact point can most easily be found by an appropriate iteration algorithm. If the surface is given by a fixed array of individual points, a local analytical approximation should be constructed in order to proceed as in the first case.

4 Machining of the reflector panels

For machining the 18 segments of the KOSMA telescope reflector we have modified one of our Moore M18 Aspheric Generators. The workpiece spindle was equipped with an Aerotech servo-motor and a Heidenhain encoder. In order to mount and machine the 37 inch diameter mirror segments, the spindle block was raised by 4 inch and the 16 inch travel of the x-slide extended by a 4 inch Dover Instrument precision air-bearing

slide. A motorized Professional Instruments 4R air-bearing spindle was mounted onto the Dover slide providing the rotational motion for the diamond tool.

The tip of the diamond tool was aligned at the centerline of the workpiece spindle. The cutting process was started at the center of the part with constant feed per revolution of the workpiece spindle, propagating along a growing spiral with constant radial velocity (except for the first 30 mm). Once the travel of the x-slide was exhausted, the process was interrupted, the Dover slide moved further out, the x-slide retracted by the same amount and the process continued. CNC data were transmitted in real time at a rate of 5 commands per second from the hard disk of a personal computer into the CNC control of the machine. Total running time of one complete cut was 12.5 hours for a segment of the inner ring and 14.5 hours for a segment of the outer ring. Machining parameters are listed below:

radius of diamond tool	$r_d = 12 \text{ mm}$
fly-cut radius	$r_c = 125 \text{ mm}$
fly-cutting frequency	$\nu = 40 \text{ Hz}$
radial velocity	$v = 18 \text{ mm/s}$
feed per revolution of workpiece spindle	$f = 0.5 \text{ mm}$
rate of surface generation	$v \cdot f = 9 \text{ mm}^2/\text{s}$
depth of cut for roughing	$< 100 \text{ } \mu\text{m}$
depth of cut for finishing	$< 20 \text{ } \mu\text{m}$

5 Results

The project was started in February 1995 after a four months planning period. By January 1996, all hard- and software modifications were completed but nothing worked, and a five months retrofitting and testing phase began. By June 1996 we finally had a safe process and finished all 18 segments within three months. The reflector was installed on schedule in September 1996.

The figure errors of the machined reflector panels were evaluated with a Ferranti coordinate measuring machine. A typical contour error map, consisting of 209 data points, is presented in **fig. 4**. The measured rms figure errors of the inner segments range from $3.1 \text{ } \mu\text{m}$ to $5.0 \text{ } \mu\text{m}$, those of the somewhat larger outer segments from $4.5 \text{ } \mu\text{m}$ to $6.4 \text{ } \mu\text{m}$. The achieved figure accuracy was not limited by the nature of the SPFC process but by the wafered design of the light-weight aluminum panels leaving a wall thickness of only 2 mm for the reflecting surface to be machined. Elastic deflection by cutting forces and distortions due to stress induced by machining were almost unavoidable.

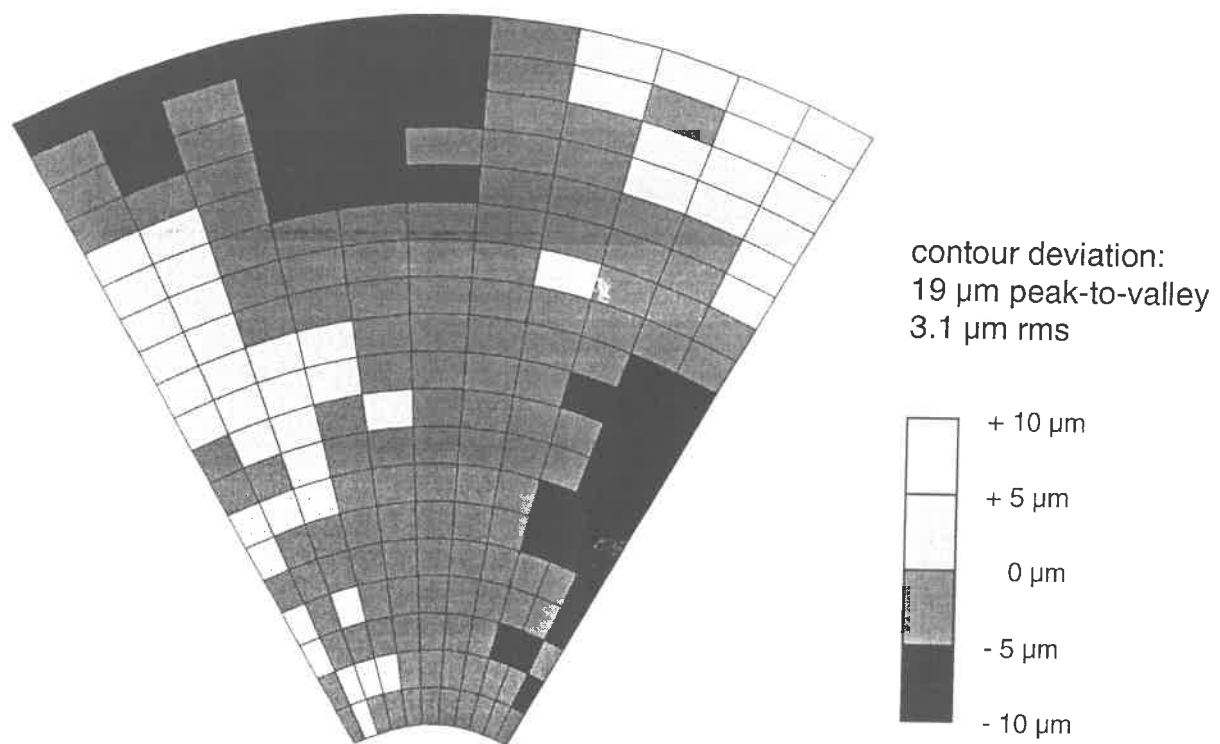


Fig. 4: Contour error map of an inner segment of the KOSMA primary reflector.

6 Acknowledgment

We deeply acknowledge the confidence and support we experienced from our contracting partner Vertex Antennentechnik GmbH who accepted the financial risk of the project. We also enjoyed fruitful discussions with the astrophysicists of the University of Cologne. This project, however, would never have succeeded without the enthusiasm of many individuals who worked hard to overcome the difficulties in one or the other critical stage of the project.

7 References

- [1] Y. Takeuchi, K. Kato, Sh. Kawakita, K. Sawada, T. Sata. „Generation of sculptured surfaces by means of an ultraprecision milling machine“. Annals of the CIRP Vol. 42/1, 1993:p. 611.
- [2] P.J. Falter, Th.A. Dow. „A diamond turning apparatus for fabrication of non-rotationally symmetric surfaces“. Ultraprecision in Manufacturing Engineering, ed. by M. Weck and R. Hartel, Springer 1983: p. 187.
- [3] M. Weck, M. Pyra. „Diamond turning of non-symmetric laser optics“. Proceedings of the ASPE 1990: p. 59.

The effect of spindle perturbations on the quality of precision grinding*

Kenneth L. Blaedel, Jared E. Bibler, and John S. Taylor

Lawrence Livermore National Laboratory
L-537, PO Box 808
Livermore, CA 94551
[blaedel@llnl.gov 510-422-0290 fax: 510-423-2419]

Introduction

In grinding, there are many perturbations to the spindle that degrade the quality of the workpiece surface being ground. Imbalance of the grinding wheel and runout of the wheel are two examples of such perturbations. These disturbances manifest themselves as varying system forces or varying system displacements. The dynamics describing the combination of machinery and material removal process then determine how these perturbations affect the process, for example, the grinding force or grinding depth of cut. The system output is given in terms of the degradation in quality of the workpiece surface. It is only in the context of the whole system, as shown in Figure 1, that the effect of poor wheel balance and poor wheel truing on workpiece quality can be assessed.

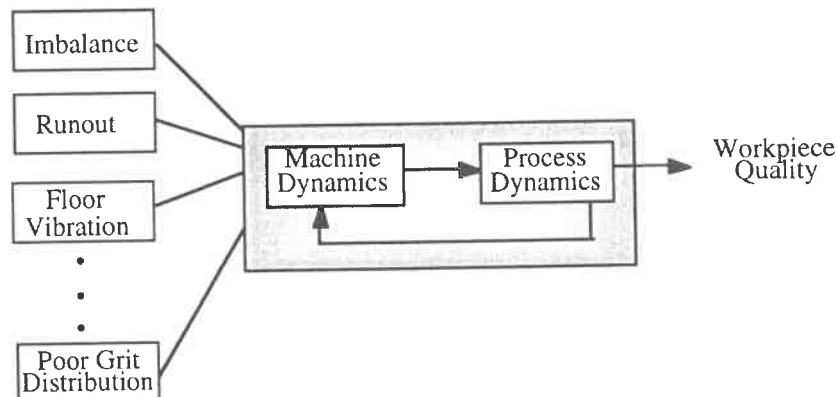


Figure 1.

Error sources (Input)

Sources of disturbances to an air-bearing spindle's axis of rotation are generally lumped into two categories, those which occur synchronously with the rotation, and those which are characterized as asynchronous. This categorization best applies to a disturbance whose frequency is at or above the fundamental frequency of the spindle, particularly in a highly repeatable air-bearing spindle. The asynchronous errors are associated with surface roughness of a ground part, and synchronous errors at the spindle rotational frequency may, depending on the relative rotational velocity of the workpiece and the wheel, manifest themselves as waviness or also as roughness in the workpiece.

In addition to the disturbances discussed above, there are subharmonic disturbances to the axis-of-rotation that usually show up as waviness, contour, or size error in the workpiece. These disturbances may be slowly varying, as for example, thermal drift of the spindle's axis, or much quicker disturbances like that induced by a change in spindle speed. In addition, for a rolling-element bearing, there are subharmonic disturbances associated with, for example, the cage pass frequency.

There are also other contributors that are not direct perturbations of the axis of rotation, but are perturbations of the material removal process and therefore indirectly of the axis of rotation. Consider, for example, that the grit is not evenly distributed around the wheel. This will have the

* This work was performed under the auspices of the U. S. Dept. of Energy by LLNL under contract No. W-7405-Eng-48.

effect of varying the normal and tangential forces associated with material removal and will consequently affect the axis of rotation depending on the machine dynamics.

These error contributors all affect the spindle axis of rotation in one way or another and they are coupled through the system dynamics model derived below. The simplified case discussed here involves the use of a high quality, air-bearing spindle where imbalance and runout become the dominant contributors to the synchronous, once-per-revolution spindle error motion. Surface finish due to asynchronous spindle error motion, and subharmonic effects will not be included in this analysis.

System Dynamics (transfer function)

In order to couple the error sources and to assess their effects on the workpiece surface, a model of the structural dynamics of the machine tool and the workpiece-tool interaction is derived with enough degrees of freedom to accurately represent the relevant parameters in the study of imbalance and runout. These parameters include the masses of the rotating spindles, the compliances between the spindles and ground, damping within the system, and the force due to infeed rate.

For cylindrical grinding, the unidirectional model illustrated in Figure 2 is driven with both an oscillating imbalance force and a runout displacement, to observe the relative effects of each. These error sources are coupled and phase-shifted to show the effects of combined imbalance and runout in the wheel. The effects of the imbalance and runout on the surface profile are observed through the material removal model. In this way, the effect of imbalance may be separated from the effect of runout as error sources.

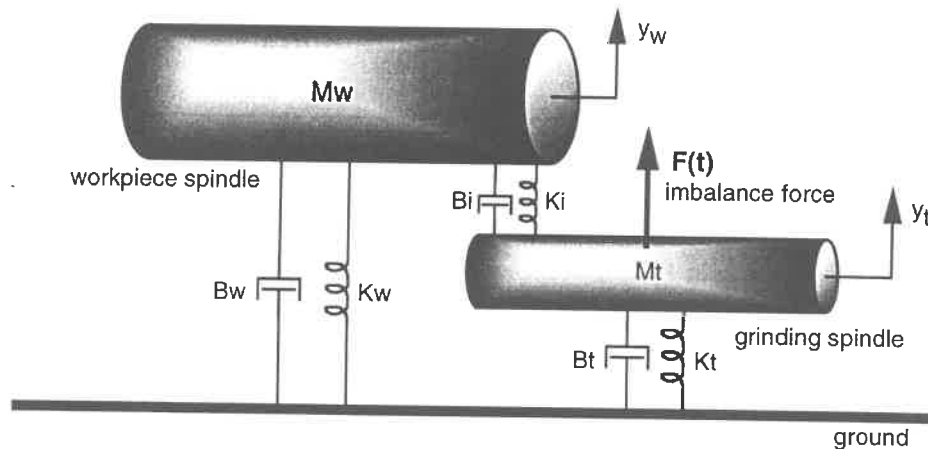


Figure 2.

In addition to investigating the effect of imbalance versus runout, the model is used to examine the importance of other system parameters with regard to imbalance in the grinding spindle. We show the effect of imbalance on cutting force while varying damping of the cutting process, that is, the damping between the grinding wheel and the workpiece. We also vary compliances in the system and observe how imbalance changes the cutting force with these new parameter values. In addition, the model simulates the synchronous loss of contact between the grinding wheel and the workpiece, reflecting, for example, a "low spot" on the grinding wheel.

The model is constructed classically, using ordinary differential equations and simulated with MATLAB and SIMULINK integration routines for stiff systems.

Workpiece Quality (Output)

The grinding performance is a combination of workpiece surface quality and cost (material removal rate, g-ratio, etc.). Surface errors are traditionally broken down into roughness, waviness, contour error, and subsurface damage. The waviness length of the once-per-revolution feed is the most sensitive category of workpiece error to imbalance and runout of the grinding wheel.

Simulation and Experimental Corroboration

The simulations performed were to address a number of fundamental questions. First, on what basis can runout and imbalance be compared, that is, what level of imbalance yields the same degradation of workpiece quality as a given level of runout? A basis for comparison is the variation in normal grinding force; when a given runout causes a grinding force variation equal to that caused by a given imbalance, their degradation of workpiece surface quality is equivalent. Analytically then, the dynamics of the interface between the wheel and workpiece determine the equivalency of imbalance and runout.

Second, given an allowable workpiece waviness, what levels of runout and imbalance are tolerable? This is determined both analytically and empirically and is similar to the approach used by Trmal and Kaliszer¹ who specified that the peak-to-valley amplitude of the waviness component introduced in the workpiece from imbalance should be less than 1/7 of the peak-to-valley height of the specified surface profile.

Third, how successfully can the deleterious effects of wheel runout be compensated by using imbalance which, at a given speed, yields an out-of-phase force in the direction normal to the workpiece surface? This compensates a sinusoidal time-varying force (due to runout) with a rotating force of constant magnitude (due to imbalance). Note that only the once-per-revolution component of runout can be compensated. This is accomplished by performing the balancing during grinding instead of in preparation for grinding. The accelerometer, which is arranged in the direction normal to the surface, will measure the sum of the accelerations due to imbalance and runout. Alternatively, this can be accomplished by using a dynamometer that measures the grinding force component normal to the surface.

Our experimental setup consists of an air spindle mounted on an isolated granite table. The spindle is imbalanced during operation and a spindle analyzer is used to determine axial and radial error motions of the spindle. A force dynamometer is mounted between the spindle and the table to measure forces generated by the imbalanced spindle in three directions. The spindle is configured for a peripheral grinding wheel and a linear table that feeds the work spindle into the grinding wheel. Surface finish on the workpiece is examined with a contacting profilometer. Compliance within the spindle is varied by changing the air supply pressure. This setup is particularly configured to explore the coupling of imbalance with a controlled compliance in the structure. This allows a "clean" experiment for corroborating the simple model of the system dynamics. Other elements of the experimental setup can also be changed, for example, by changing the wheel and workpiece materials, but such a change has a far ranging effect on many of the system parameters.

Conclusions

Considering perturbations to the grinding spindle along with the machine and process dynamics allows the effects of imbalance to be directly compared to wheel runout in terms of degradation of the workpiece surface caused by each. In addition, this leads to a rational method for specifying the quality of balance and runout that are required for a particular grinding operation to achieve a given workpiece surface quality. A different performance index is proposed to achieve higher quality grinding, that of minimizing the force perturbations to the grinding process rather than the force perturbations to the machine. This is tantamount to partially compensating once-per-revolution wheel runout with wheel imbalance.

¹ G Trmal and H Kaliszer, "Some Aspects of Unbalance in the Wheel-Spindle Assembly of Cylindrical Grinding Machines," *Annals of the CIRP*, 21/1/1977, pp89-90.

POSSIBLE MECHANISM OF CRACK INITIATION IN MICROMACHINING OF SILICON

Shoichi Shimada

Osaka University, Suita, Osaka, Japan

Toyoshiro Inamura, Nobuhiro Takezawa

Nagoya Institute of Technology, Nagoya, Japan

Naoya Ikawa

Osaka Electro-Communication University, Neyagawa, Osaka, Japan

Yoshiaki Onchi

Mizuho Co., Ltd., Joyo, Kyoto, Japan

Toshio Sata

Toyota Technological Institute, Nagoya, Japan

BACKGROUND

With the improvement of ultraprecision machine tools, which have smaller error motion and higher stiffness, ductile mode machining of brittle materials has become an attractive technique to achieve higher form accuracy, better surface finish and lower costs. Single point diamond turning is one of the most useful technique for the machining of variety of crystalline and infrared materials.

It is well known that there is a transition in material removal mechanism of brittle materials from brittle to ductile under the depth of cut smaller than a critical value which intrinsic to the workmaterial to be machined[1-5]. The explanation of transition mechanism based on the energy balance in different removal processes can be qualitatively understood well. However, attention should be paid to a premise, in which the pre-existing cracks are assumed, of the hypothesis. In advanced crystalline or semiconductor materials, defect or microcrack which are corresponding to a pre-existing Griffith crack can scarcely observed. Therefore, the transition mechanism in the machining of defect-free brittle materials can not be explained by the hypothesis described above.

In this paper, to understand the mechanisms of ductile mode material removal at extremely small depth of cut and the transition in material removal from brittle to ductile in machining of defect-free brittle material, a cutting simulation of monocrystalline Si are carried out based on molecular dynamics (MD).

MOLECULAR DYNAMICS SIMULATION

MD simulations are carried out using a three dimensional model shown in Fig.1. A cubic monocrystalline Si workpiece which has (111) top surface and (110) side surfaces with periodic boundaries is cut in $\langle 112 \rangle$ direction by diamond cutting edge with 2 nm edge radius. The thickness of unit cell of the model are 1.5 nm (8 atomic layers). Both of the tools and workpiece consist of fixed boundary atoms shown by dense shadow, thermostat atoms shown by medium shadow and Newtonian atoms shown by open circle. In the thermostat layer, the equivalent temperature of the atoms are adjusted for every computation time steps so as to maintain the average temperature in the layer at 300 K. Tersoff type three body potentials are used for the interatomic potential between Si-Si, C-C and Si-C. The uncut chip thickness and cutting speed are set to be 1 nm and 200 m/s, respectively. Although this cutting speed is unrealistic, it is employed to reduce computation time after the confirmation that there is no remarkable difference between the deformation behavior in the cutting under 200 m/s and 20 m/s.

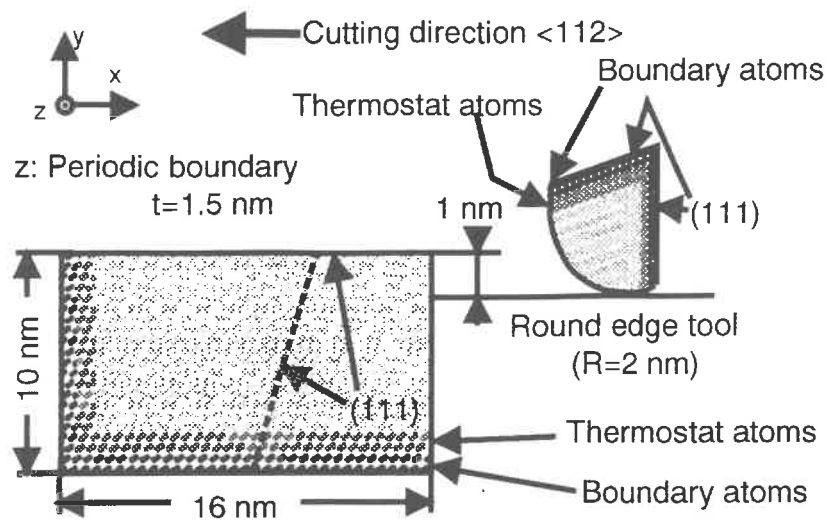


Fig.1 Schematic model of microcutting of Si for MD simulation

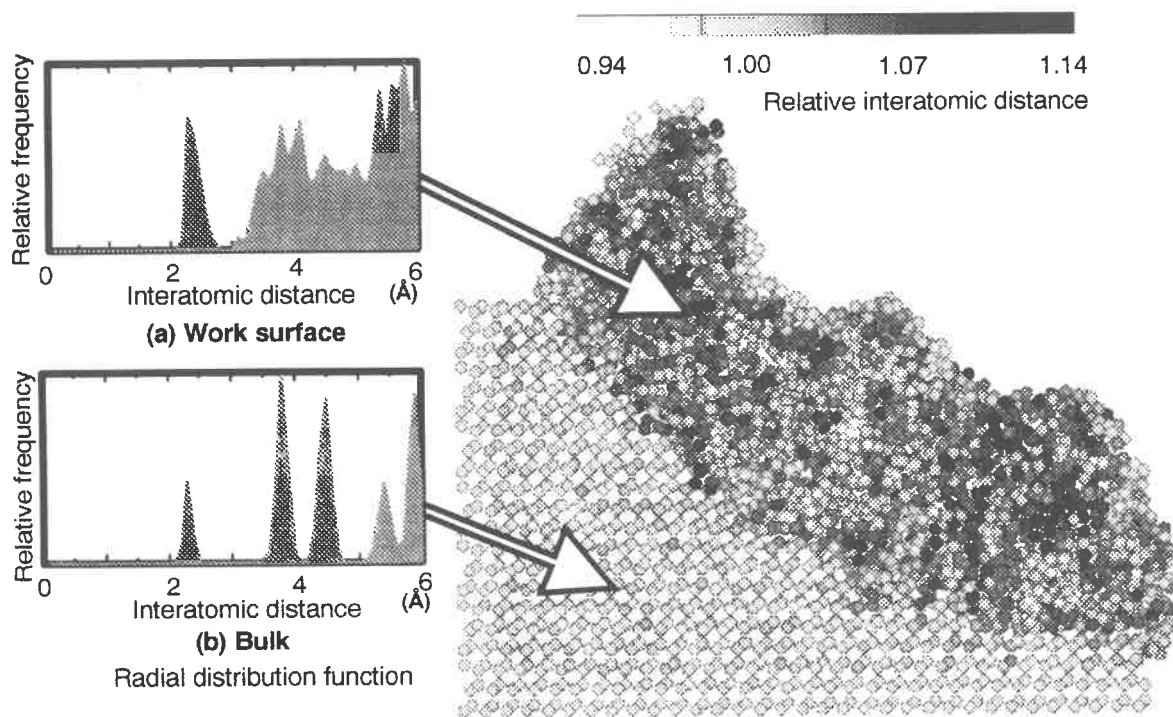


Fig.2 Structure of microturned surface of Si

MICROSTRUCTURE OF SURFACE MACHINED

Figure 2 shows a snapshot of the model in the cutting simulation. Although deformed layer is observed in the workpiece around the cutting edge and the work surface machined, no fracture or cleavage can be observed. The material removal seems to take place in ductile mode under such a small depth of cut. The radial distribution function of interatomic distance on the work surface shows characteristic feature for an amorphous structure as shown in Fig.2. As the average equivalent temperature of the atoms in deformed layer rises to about 3000 K, the structure of the layer changes to amorphous state by rapid cooling after the passage of cutting edge. On the other hand, Fig.3 shows a TEM micrograph of a cross section of the work surface machined in ductile mode by diamond turning under the uncut chip thickness of 30 nm. The work surface is covered by a thin layer the thickness of which is about 200 nm. This layer is confirmed to have amorphous structure from electron diffraction pattern. Figure 2 also shows the distribution of relative interatomic distance in cooled work surface to the average value in strain free bulk Si at 300 K. The interatomic distance in the surface layer is expanded due to phase transition into amorphous state. A few percent extension can also be observed in the crystalline bulk near the surface layer. The result suggests that residual tensile strain remains in crystalline bulk below the surface amorphous layer and that the larger residual strain, which is predicted to remain under larger uncut chip thickness, may induce some sub-surface damages.

TRANSITION IN MATERIAL REMOVAL MECHANISM

Figure 4 shows a snapshot corresponding to Fig.2 in a sliced section with the thickness of two atomic layers. A kind of cavities or microcrevices, in which the atoms are very sparsely distributed, are observed in the deformed layer due to the violent vibration of workpiece atoms under elevated temperature. Successive observations of the snapshots shows that the microcrevices are instantaneously

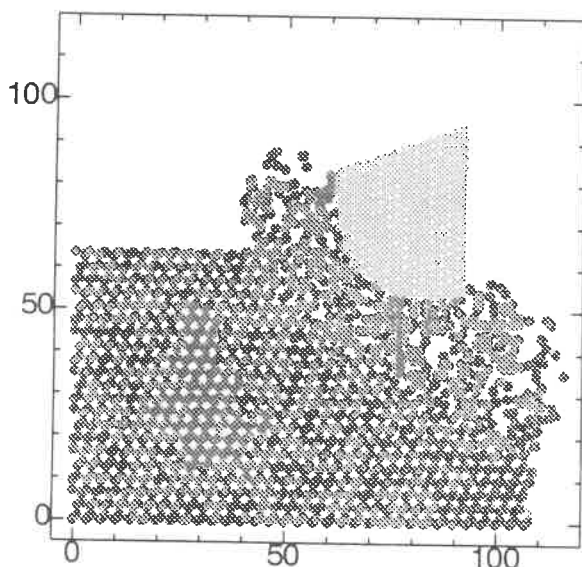


Fig.3 A snapshot of sliced section of workpiece and cutting tool in microcutting of Si

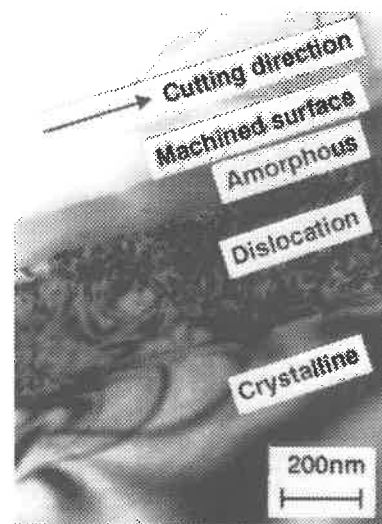


Fig.4 TEM micrograph of cross section of Si work surface machined by diamond turning in ductile mode

induced by the propagation and interference of shock waves which are repeatedly generated in cutting zone when the lattice of workpiece is crushed by the plowing of cutting edge. Cutting simulation under the uncut chip thickness larger than the critical value cannot be carried out yet due to the limitation in computation. However, the width and length of the microcrevices are considered to be governed by the amount of energy of elastic waves. The energy of the elastic waves is supplied from the potential energy stored in the workpiece around the cutting edge. Therefore, the energy released may have the amount under larger uncut chip thickness enough to create a microcrack on the work surface or to remain misfit in lattice plane due to the prevention of perfect rearrangement when the microcrevices are closed. Furthermore, atmospheric molecules may have a chance to invade into the microcrevices. The adsorption of the molecules on the inside surface of the microcrevices causes the decrease of surface energy of the workpiece. In such cases, brittle mode material removal may take place even in machining of a defect-free brittle material.

CONCLUSIONS

To understand the mechanisms of ductile mode material removal at extremely small depth of cut and brittle-ductile transition in material removal process, microcutting simulation of defect-free monocrystalline Si are carried out based on molecular dynamics. The MD simulation shows that the surface layer on work surface changes to amorphous structure. As the larger tensile residual strain remains on the work surface, sub-surface damage may be induced when machined under the larger uncut chip thickness. The simulation also suggests that the critical depth of cut can be governed by the amount of energy of elastic waves generated in cutting zone as the results of release of potential energy stored in workpiece by the plowing of cutting edge. Therefore, in case of large scale cutting, brittle mode material removal may take place even on a defect-free brittle material.

REFERENCES

- [1] Bifano, T.G., Dow, T.G., Scattergood, R.O., 1991, Ductile-Regime Grinding : A New Technology for Machining Brittle Materials, Trans. ASME : J. Engg. for Industry, 113, pp.184-489.
- [2] Blakley, W.S., Scattergood, R.O., 1991, Ductile Regime Machining Model for Diamond Turning of Brittle Materials, Precision Engineering, 13, 2, pp.95 -103.
- [3] Nakasuji, T., Koderu, S., Hara, S., Matsunaga, H., Ikawa, N., Shimada, S., 1990, Diamond Turning of Brittle Materials for Optical Components, Annals of the CIRP, 39, 1, pp.89-92.
- [4] Puttick, K.E., 1980, The Correlation of Fracture Transitions, J. Phys., D: Appl. Phys., 13, pp.2249-2262.
- [5] Puttick, K.E., Rudman, M.S., Smith, K.J., Franks, A., 1989, Single -Point Diamond Machining of Glasses, Proc. R. Soc. Lond., A426, pp.19-30.

PRECISION MANUFACTURING OF COMPACT DISC MASTER STAMPERS

Thomas Bifano and Helen Caggiano, *Boston University, College of Engineering*
Paul Bierden, *Prism Corporation*

ABSTRACT

We have developed a new manufacturing process for compact disc master stamper production. *Stamper* is an industry term for the mold used to replicate polymer CDs in an injection molding device. The stamper surface contains a negative image of the CD's >1 billion features, each measuring 0.5 μm wide, 1-3 μm long, and 150 nm tall. These features are embossed in a spiral pattern on the 138 mm diameter stamper substrate. In conventional mastering, nickel sub-master stampers are generated by electroforming from a glass master with a masked photoresist layer. In the new process, called Precision Stamper Manufacturing (PSM), sub-master stampers are not required – the ceramic master is used directly as a stamper for injection molding. The technology uses precision manufacturing to improve productivity, reduce hazardous waste, and reduce costs in production of master stampers. All critical process steps have been proven feasible in the research reported here. More than 120 minutes of conventional manufacturing processes are replaced by about 20 minutes of precision machining processes, with improved precision in the final component. Another benefit of PSM is that it eliminates most process steps that produce hazardous wastes.

Objectives of the experimental research have been to verify feasibility of the method, and to manufacture a prototype CD stamper. Over the past year, we have demonstrated feasibility of all process steps in the new method, and results will be presented. Also, we uncovered and quantified several important limitations for each process step, and combined the individual steps into a manufacturing cycle through which a full-scale working stamper prototype was made. These experiments made use of available facilities at two major equipment suppliers for the CD, an ion machining equipment, and a national research facility, in addition to facilities at Boston University and Prism Corporation.

MANUFACTURING CD STAMPERS

Major steps in the conventional process for fabricating CD stampers are compared with those of the newly proposed process in Figure 1¹. In the conventional process, depicted on the right, a flat, polished glass substrate disk is cleaned and coated with positive-tone photoresist. Next a focused laser beam traces a spiral path on the disc surface. The laser is switched on and off according to a streaming binary data signal. After recording, the photoresist is developed, leaving elongated pits where it was irradiated. The patterned photoresist is chemically metalized, and a thick film of nickel is grown on top of this layer by electroforming. This nickel *sub-master stamper* is peeled away from the master disc substrate, cleaned, and protected. The stamper back-side is polished, and its inner and outer diameters are punched on a press. The protective coating is removed, leaving the stamper ready for use in an injection molding machine to replicate discs. The total manufacturing time to make a single CD stamper is a little more than 4 hours.

The PSM process is illustrated schematically on the left side of Figure 1. It calls for two major changes in the manufacturing process for stampers:

- replacement of the glass master disk substrate by a tough ceramic, which will be used as the injection mold stamper.
- replacement of electroforming, peeling, protecting, polishing, and punching operations by a single ion machining step.

¹Currently, there are at least three major vendors of OEM master stamper fabrication systems. The processes described here are specific to one product of the leading vendor. Other systems' processes vary slightly in detail.

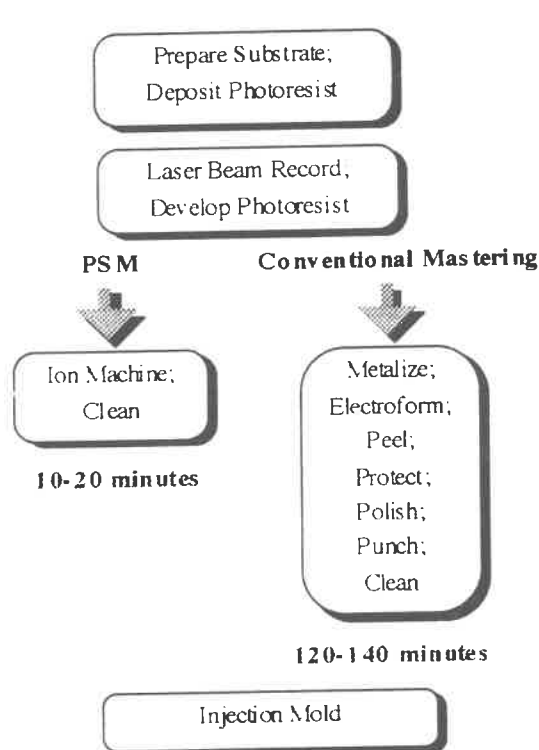


Figure 1: Flow-Chart comparison between PSM and conventional stamper manufacturing

forming are the sources of nearly all of the six liters of toxic and hazardous waste that is generated in producing a stamper.

Ion machining as a precision material removal process has been refined for use in sub-aperture shaping of glass and ceramic substrates, and for various high technology applications such as fabrication of individual three-dimensional microstructures, sharpening of scanning probe tips, and imprinting patterns on carbides^{2,3,4,5}. The process, which is distinct from the more well known process of *reactive ion etching*, employs a plasma-generating source or "gun" that ionizes argon gas in an evacuated chamber. While still in the source, ionized atoms (usually argon, but helium and other inert gasses have also been used) are accelerated by a DC electric field through a grid-shaped aperture. As they leave the source at high velocity, the collimated beam of ions is neutralized by an oblique beam of electrons emerging from an adjacent source. The stream of argon molecules, now chemically and electrically inert, impact a target surface and sputter molecules from that surface in a finely controlled erosion process. Typical beam current densities are 1-2 mA cm², accelerated by a 1000 volt potential. For many ceramic and glass materials, this results in a sputtering rate of several tens of nanometers per minute. Sources range in diameter from 30 mm to more than 600 mm, and a number of unusual geometries of aperture (and consequently beam shape) have been reported. Adaptation of ion machining to optical surface contouring was first demonstrated by McNeil, et. al.⁶ The most visible suc-

All other processes, including cleaning the master disk, spinning-on a photoresist layer, laser recording, and photoresist development will remain largely unchanged. That is, the process is compatible with the equipment that immediately precedes and follows stamper fabrication (e.g. laser recording systems and injection molding systems). In PSM, production of a sub-master nickel stamper is no longer necessary. The ion machined ceramic master stamper is used *directly* for polymer injection molding. The overall processing time (including laser recording) is reduced by a factor of 2 (from more than four hours to about two hours) from the conventional manufacturing process. More than 120 minutes of the conventional manufacturing processes are replaced by a 10-20 minute ion machining process (inclusive of time for part handling and transfer through a vacuum interlock). Capital equipment costs and direct manufacturing costs for the proposed process are lower than that required for the present process. Other benefits of the proposed process in comparison with the conventional process are that it can be automated, and it eliminates all process steps that produce toxic and hazardous wastes. Metalization and electro-

² Fawcett, S. C., Bifano, T. G., and Drueding, T., "Neutral Ion Figuring of Chemically Vapor Deposited Silicon Carbide," *Optical Engineering*, [33]3, pp. 967-974, 1994.

³ Miyamoto, I., and Shuhara, A., "Ion Beam Machining of Tungsten Carbide Chips - Fabrication of Fine Patterns," *Annals of the CIRP*, Vol. 40, No. 1, 1991.

⁴ Miyamoto, I., Ezawa, T., and Itabashi, K., "Ion Beam Fabrication of Diamond Probes for a Scanning Tunneling Microscope," *Nanotechnology*, Vol. 2, pp. 52-56, 1991.

⁵ Egert, C.M., "Roughness Evolution of Optical Materials Induced by Ion Beam Milling," *SPIE* 1752, 1992.

⁶ McNeil, J.R. and Herrmann, Jr., W.C., "Ion Beam Applications for Precision Infrared Optics," *Journal Vacuum Science and Technology*, 20 (3), pp. 324-326, March 1982.

cess of this process came with the Keck mirror telescope project⁷. Although it was not part of the original manufacturing proposal for that project, ion machining succeeded where fine polishing had failed as a final contouring process for the Keck's 36 hexagonal optical segments (each measuring more than 1 meter across.) More recently, ion machining has been used to fabricate ceramic optical elements for space-based applications^{8,9,10,11}. The process has not been used previously in an advanced manufacturing application requiring micrometer-sized lateral features on a large surface area. Producing billions of such features with uniform good quality on a single substrate by broad beam ion machining through a mask is a principal focus of the research reported in this paper.

SUBSTRATES FOR CERAMIC STAMPER PROTOTYPES

The principal attributes required of a PSM substrate are toughness, thermal shock resistance¹², ion machinability, thermal conductivity and hardness. These properties are important for raw substrate fabrication, ion machining, and injection molding. Table 1 shows some of these properties, tabulated from various reference sources^{13,14} for sample substrate materials that were considered in our feasibility research. For PSM, it is desired that the stamper substrate has relatively high values of hardness, toughness, and thermal shock resistance. Chemical vapor deposited silicon carbide (CVD SiC) was the clear choice for this research based on its physical properties, and was chosen as the first substrate material to be evaluated

Substrate	Knoop Hardness (kg/mm ²)	Fracture Toughness (MPa m ^{1/2})	Thermal Shock Resistance (W/m)	Thermal Conductivity (W/m K)	Ion Machinability (Surface Quality)
Nickel	100	~100	7138	80	Poor
Al ₂ O ₃	2100	4	3225	30	Very Good
SiC	2500	3	19149	90	Excellent
Fused Silica	500	0.7	5867	1.3	Excellent
Glassy Carbon®	500	2	135517	120	Excellent
Corning (VGJ)®	450	1.3	546	2.5	Excellent

Table 1: Prospective stamper substrate materials and relevant properties

CD stampers are nominally 138.00 mm in diameter with a 35.40 mm diameter central hole, and 0.300 mm thick. Required finish quality is specular with roughness lower than 10 nm. SiC ceramic stamper substrates were designed to meet these specifications, and were further constrained in flatness ($\pm 5 \mu\text{m}$), roughness (1.0 nm Ra), and surface damage (none visible with 1000X optical microscopy).

⁷Allen, L.N., Keim, R.E., Lewis, T.S., and Ullom, J., "Surface Error Correction of a Keck 10m Telescope Primary Mirror Segment by Ion Figuring," *SPIE*, Vol. 1531, July 1991.

⁸Bifano, T. G., Kahl, W. K., and Yi, Y., "Fixed-Abrasive Grinding CVD Silicon Carbide Mirrors," *J. Precision Eng'g*, [16]2, pp. 109-116, 1994.

⁹Drueding, T. W., Wilson, S., Fawcett, S. C., and Bifano, T. G., "Ion Beam Figuring of Small Optical Components," *Optical Engineering*, [34]12, pp. 3565-3571, 1995.

¹⁰Drueding, T., Bifano, T. G., and Fawcett, S. C., "Contouring Algorithm for Ion Figuring," *J. Precision Eng'g*, [17]1, pp. 10-21, 1995.

¹¹Fawcett, S. C., Bifano, T. G., and Drueding, T., "Neutral Ion Figuring of Chemically Vapor Deposited Silicon Carbide," *Optical Engineering*, [33]3, pp. 967-974, 1994.

¹²Thermal shock resistance is the product of a material's modulus of rupture or tensile strength and its thermal conductivity, divided by the product of its coefficient of thermal expansion and its modulus of elasticity. Larger numbers imply greater ability to absorb rapid heating or cooling without fracture or yielding.

¹³Callister, W., *Material Science and Engineering*, John Wiley & Sons, Inc., New York, NY, 1994.

¹⁴Ashby, M., Jones, D., *Engineering Materials 1*, Pergamon Press, Elmsford, NY, 1980.

Five substrate samples were made by double sided lapping and polishing CVD SiC. The surface flatness ($\pm 1 \mu\text{m}$), measured using a commercial inteferometer, exceeded the specifications for both samples. Surface roughness, measured using a commercial AFM was 0.25 nm Ra, again better than the required specifications by a considerable margin. No surface damage was visible on the disc, using optical microscopy at 1000X.

ION MACHINING EXPERIMENTS

We assembled a working ion machining system and used it to perform a series of machining experiments. A photograph of the system is shown in Figure 2. The heart of the system is an 11 cm broad-beam ion source suspended above a rotating sample stage.

Beam voltage	1000 V
Beam current	300 mA
Argon flow rate	4.0 sccm
Base pressure	1 μTorr
Process pressure	2 mTorr

Table 2: Typical process conditions for ion machining

Samples were spin coated with a 1 μm thick negative tone photoresist and recorded on a commercially available laser beam recording system. The photoresist was baked and developed, resulting in a patterned spiral mask containing ~ 1 billion CD features. These samples were then ion machined for approximately 7 minutes to achieve an etching depth of 150 nm on the unmasked regions of the substrate.

Figure 3 shows SEM photographs comparing CD features produced in Ni using the conventional electroforming process (left) to those produced in SiC by ion machining (right). For scale, the features shown in both photos are 500 nm in width, separated by a track pitch of 1.6 μm .



Figure 2: A photograph of the ion machining system used in these experiments

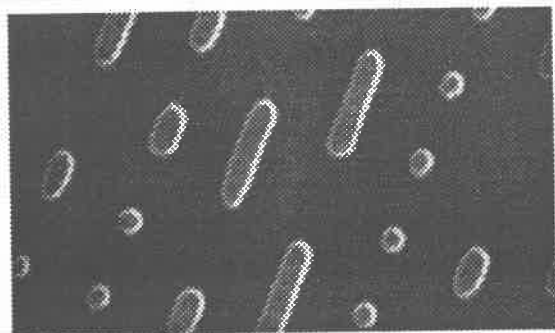
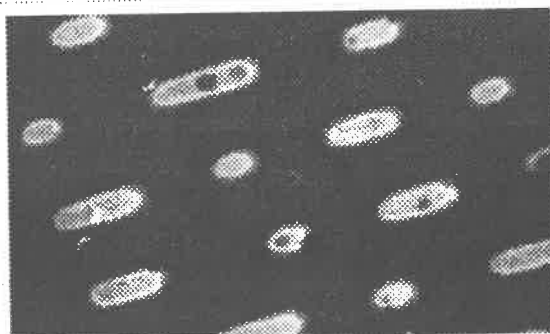


Figure 3: SEM photos of electroformed nickel (left) and ion machined ceramic (right) stampers.

INJECTION MOLDING WITH A CERAMIC STAMPER PROTOTYPE

Mold temperature	85°C
Polymer temperature	330°C
Clamping pressure	30 tons
Injection time	1 second
Cooling time	3 seconds

Table 3: Conditions for injection molding

An injection molding test was performed on a commercial CD molding machine. Figure 4 is a photo of the first of several CD replicas that were molded using a ceramic ion machined stamper. Molding took place using the conditions outlined in Table 3. Work is continuing on this part of the research, to quantify stamper performance in terms of CD quality and stamper lifetime as a function of molding parameters.

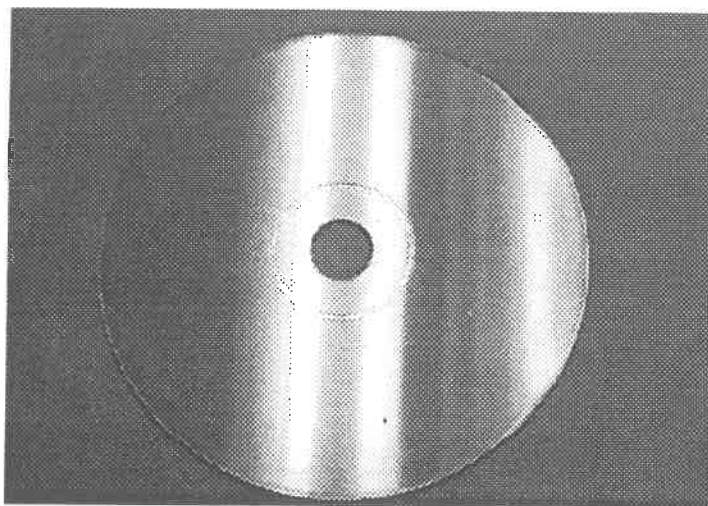


Figure 4: Photograph of the first CD replica injection molded with an ion-machined SiC stamper

ANALYSIS OF STAMPER QUALITY AND FEATURE GEOMETRY

An atomic force microscope profile of a typical feature on an ion machined stamper prototype is shown in Figure 5. Several important observations can be made from this data. First, it is clear that both the

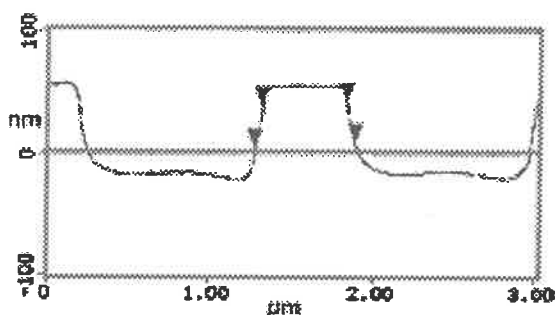


Figure 5: Atomic force microscope profile of an ion machined feature

masked areas (on top of the bump) and the unmasked areas have similar roughness. That is, the roughness of the sample was not increased by ion machining. Second, the wall slope of the feature is seen to be $\sim 35^\circ$. The observed level of wall slope is, in fact, ideally suited to the task of replicating CDs through injection molding. Substantially steeper slopes would adversely affect polymer release from the mold, while shallower slopes would reduce the signal level produced by the CD in a player. No feature analysis on polymer replicas has been performed to date.

CONCLUSIONS

A novel process for precision manufacture of optical disc stampers has been introduced in this paper. The process makes use of ceramic instead of nickel substrates, and uses neutral ion machining instead of electroforming as a fabrication method. It promises to be faster, cleaner, more reliable, and more precise than the conventional process. Results from full-scale prototype fabrication experiments conducted on commercial equipment have verified that all production steps in the new process are feasible, and that the process is compatible with the conventional production systems that precede and follow stamper fabrication (i.e., laser beam recording and injection molding, respectively). Technical challenges related to integration and optimization of process steps have been identified, as have alternate substrate and mask materials.

ACKNOWLEDGMENTS

We are grateful for support to Boston University through NSF Grant DMI9500162, and to Prism Corporation through DOE STTR Contract DE-FG02-95ER86033.

A Tabletop Factory to Fabricate Micro Machines, "Nano Manufacturing World"

Masayuki Nakao and Yotaro Hatamura
Dept. of Engineering Synthesis, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Abstract

To fabricate micro machines, the authors designed, constructed and evaluated a tabletop factory called "Nano Manufacturing World (NMW)". We designed a compact structure for an integrated process and a vision-aided tele-operation system between the process site and the operating site. Using various working processes developed in NMW, we demonstrated fabricating some micro structures. Through the demonstration, we evaluate that NMW functions effectively as designed, and will become a solution for micro manufacturing.

Keywords: Micro Machines, Integrated Process, Manufacturing System Design

1. Introduction

In the near future, micro machines are expected to be employed in various applications; information processing devices, medical surgery equipment and science instruments.

Generally, micro machines have 3D structures whose typical dimensions range from 10 μm to 1 mm. A micro machining process such as grinding, laser cutting or electro-discharge machining forms the overall shape. Meanwhile, micro machines engage micro 2.5D functional features on the original 3D structure. These features are typically 1 to 100 nm thick giving them the 2.5D appearance. A semi-conductor production process such as molecular beam deposition, sputtering deposition, heat oxidation or ion implantation forms these film features.

An example is the magnetic head of a HDD (Hard Disk Drive). The overall shape is a rectangular structure ground to some 100 μm cube. The head is also equipped with film features on it: a multi-layer magnetic circuit, a low-friction sliding surface or a tracking fine actuator. These features are approximately 10 to 100 nm thick.

In producing micro machines, integration of the above two processes is indispensable [1]. Many researchers have developed individual processes, but there has been few work on the integration. A series connection of the processes will produce a huge manufacturing system.

In designing such a micro manufacturing system, there are two large obstacles: (1) a micro machine can not be fabricated by a micro machine and (2) a micro machine is made on a black box.

Nowadays, semi-conductor researchers or physicists develop many processes for ultra-thin films. A typical process, however, consists of large ($> 8 \text{ m}^3$) fabrication machines interconnected with long ($> 10 \text{ m}$) vacuum tunnels. Moreover, the thin film formation inside a vacuum chamber cannot be observed. Attempts have been made using a scanning electron microscope (SEM) or an atomic force microscope (AFM). Such equipment are too large to install inside fabrication machines.

To overcome the obstacles, the authors designed and constructed a tabletop factory called "Nano Manufacturing World (NMW)". We shrunk a large manufacturing system into the

micro fabrication site and opened a tele-operation interface to the operation site.

2. Design of NMW

A fabrication machine needs two technical issues addressed: the working process it employs and the total structure synthesizing its components.

First, we developed an integrated working process following a conventional factory setting. NMW forms and assembles a 3D structure, cuts micro 2.5D features on the 3D structure, and finally conducts post-machining processes such as inspection, adjustment or trial-run.

To fulfill these tasks, we performed and evaluated many elementary processes: tool machining, dry etching, fixturing, handling, molding, soldering, etc. We also developed many new tools to meet various functional needs. For example, FAB (Fast Atom Beam) etching technology we developed fabricates a micro 2.5D feature as Fig. 1 shows. The thin plate array in the figure has vertical side walls

and smoothly etched bottom floors [2]. FAB etching can machine nanometer-sized structures such as quantum elements or micro multi-lens [3].

Secondly, we designed a total NMW structure to overcome the obstacles.

We could compact the structure with the strategy of a "multi-task on one table". Fig. 2 illustrates the total NMW structure. It has two

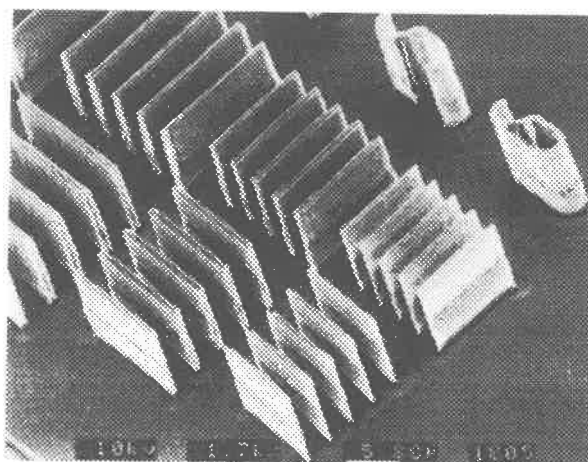


Fig. 1 SEM view of a thin plate array etched by Fast Atom Beam (Plates: Ga-As, FAB: Chlorine)

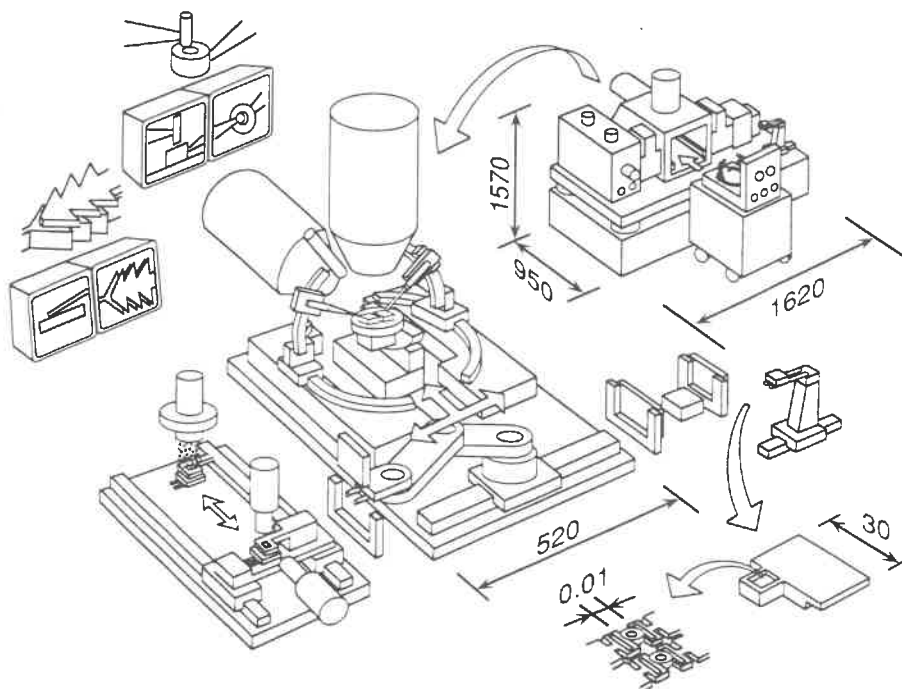


Fig. 2 Schematic view of the total structure of NMW

tables: the shape-making table (left) and the handling table (right). Around the handling table, a Concentrated Motion Manipulator (CMM) and a Multi-Directional SEM (MD-SEM) are installed.

In a conventional factory setting, we move workpieces among fixed fabrication machines. NMW adversely moves fabrication machines onto the fixed workpiece to avoid the need for accurate re-fixturing or searching for lost pieces.

The structure we designed, however, was larger than we wanted. The limiting factor is the size of components even though the workpiece itself can be designed smaller from 100 mm to 10 μm . For example, the smallest magneto-electric motor is 1 mm or greater; the smallest SEM gun is larger than a CCD camera. Even for moving a 10 μm part, the manipulator must move it to the next pallet 10 mm away.

NMW has a special tele-operating function which conventional factories do not have; an interface between the micro physical phenomena and the macro operational environment. For a human operator to make proper judgment and gain knowledge about the processes, dynamic, realtime information (vision, force feedback, or sound) from the micro working site is needed through microscopes or various sensors.

3. Construction of NMW

Next we constructed the NMW as designed. Fig. 3 shows its appearance. The approximate dimensions are 1.6 m long, 0.9 m wide, and 1.6 m tall. It has 2 major vacuum rooms mounted on a magnetically controlled damping table. One is for dry etching, sputtering and diffusion connection using FAB. The other is for tool machining, assembly and inspection using various tools attached to the right and left hand of CMM. CMM moves about a sta-

tionary point which is the observation point of MD-SEM.

CMM has a table and two hands with a total of 20 degrees of freedom in translation and rotation [4]. CMM has all the 8 rotational centers within the focused area of the microscopes using circular guides, so that tool tips and workpieces are always observed. MD-SEM has two electron guns, one on the top and the other on the side. We had developed a 3D-SEM where two inclined beams scanned an observed area alternately [5], but now verified that MD-SEM is easier to judge the distance between a tool and a workpiece. MD-SEM is effective in preventing tool and workpiece collision due to the side view and the shadow of the tool on the workpiece. Even a regular SEM can show the shadow because the tool obstructs an electron reflected against the workpiece.

The chambers are maintained at 10^{-6} torr in vacuum and less than 1 nm p-p in vibration even when stepping motors or turbo vacuum pumps are in operation. An R- θ robot transfers a pallet (29 x 25 x 6 mm); the pallet fixes a micro workpiece by a spring; a vibrating cutting tool or a pencil-typed FAB source machines the workpiece [2]; an electrostatic handling tool moves the workpiece [5].

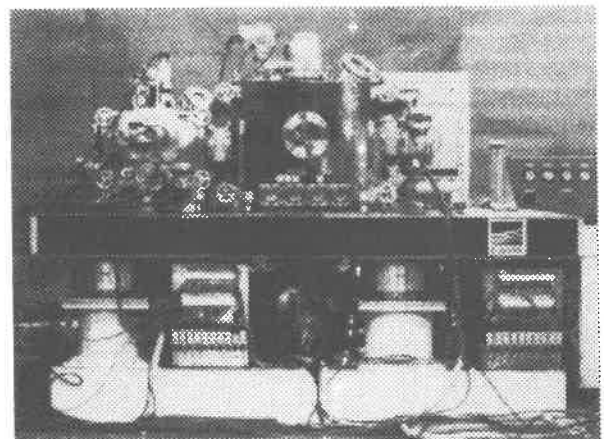


Fig. 3 Nano Manufacturing World (Length of the table: 1.6 m)

To accomplish an actual working NMW, we incorporated various components to it: a miniature slide bearing, a micro force sensor, a ultra-fine positioner [3], a thermal deformation sensor and so on. For example, the ultra-fine positioner, driven by piezo elements and made of Invar alloy, moves with 1 nm resolution.

Fig. 4 depicts a tele-operation system called "Image Driven Operation System (IDOS)". We actively used microscopic images. The image served vision-feedback control of CMM. To save a space, the manipulators in NMW do not have any encoders for position-feedback control. The vision-feedback control was realized by realtime, high-speed processors for vision tracking. The images were also invoked for memorization or repetition. Most of all, vision helps better operatability by transferring real operation feeling.

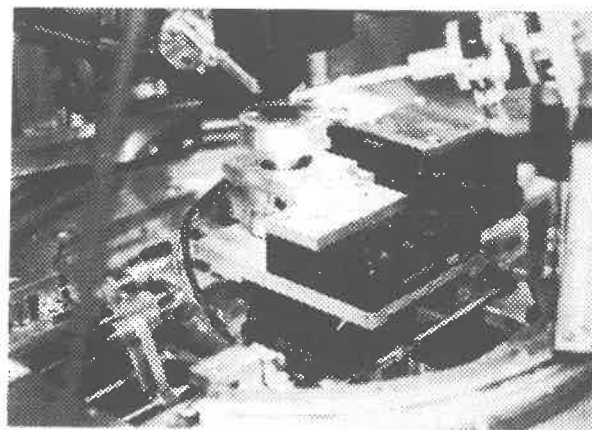


Fig. 5 Operating the Concentrated Motion Manipulator to fold an "origami" airplane (Diameter of circular guide: 300 mm)

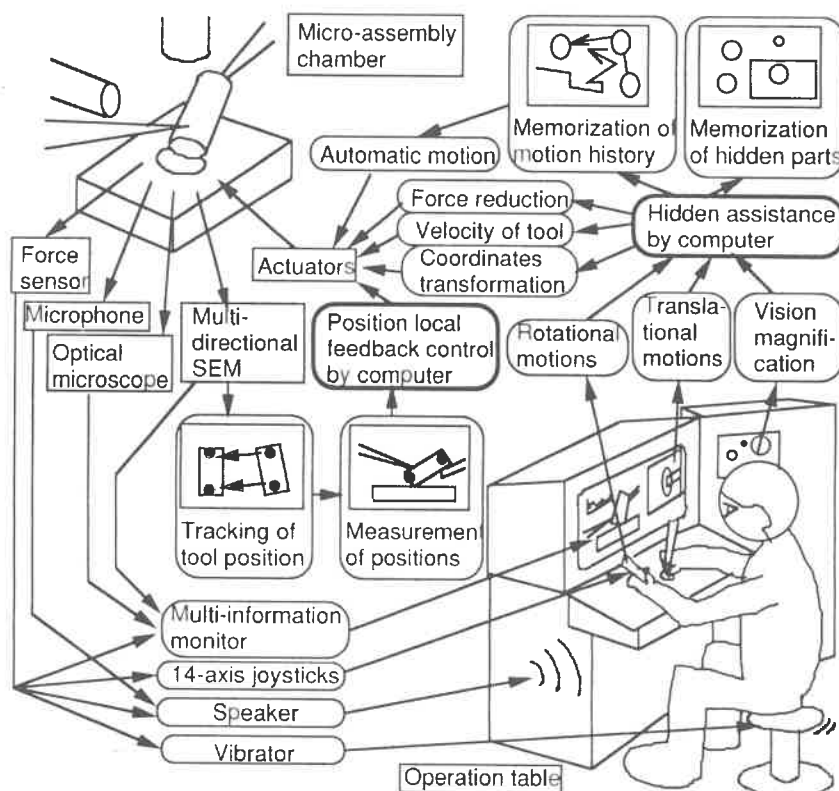


Fig. 4 Schematic view of Image Driven Operation System (IDOS)

thickness is indispensable for those technologies.

It has been confirmed through the fabrications that each mechanism employed in NMW functions as designed, and that the physical phenomena encountered in NMW are effectively transmitted to the operator through the developed interface, IDOS.

5. Conclusion

To fabricate micro machines, the authors designed, constructed and evaluated a table-top factory called "Nano Manufacturing World (NMW)". We designed a compact structure for an integrated working process and image-aided tele-operation system between the process site and the operating site called "Image Driven Operation System (IDOS)". Using various specific processes for micro structures, we actually fabricated some 3D micro structures. Through the demonstration, even at its relatively early stage, we revealed that the design of NMW functions effectively, and NMW will become a solution for micro manufacturing systems.

The authors would like to thank Dr. Masahiro Hatakeyama, Mr. Katsunori Ichiki, Mr. Masaaki Kajiyama and Mr. Takashi Tsuzuki of Ebara Research Co., Ltd. for improving FAB etching, the magnetically controlled damping table and the R- θ robot, and Prof. Tomomasa Sato and Mr. Koichi Koyano of the University of Tokyo for their help in developing of CMM.

Reference

- [1] Y. Hatamura, et al.; Construction of Nano Manufacturing World, *Microsystems Technologies*, 1-3, 155, 1995
- [2] M. Nakao, et al.; Integration of 3-D Shape Construction and Assembly to Realize Micro Systems, *Microsystems Technologies*, 1-3,

124, 1995

[3] M. Nakao, et al.; Breakthroughs in 3-D Ultra-fine Shape Fabrication by Dynamic FAB Etching, *Proc. of ASPE '95*, 376, 1995

[4] M. Nakao, et al.; Structure and Working Characteristic of Concentrated Motion Manipulator in Nano Manufacturing World, *Proc. of JSME*, 95-1(iv), 277, 1995

[5] M. Nakao, et al.; Micro Assembly in Vacuum under SEM, *Proc. of JSME*, 95-1(iv), 285, 1995

In-Situ Slurry Film Measurements for Chemical-Mechanical Polishing

J. Levert (1), R. Baker (2), F. Mess, (1) S. Danyluk(1), R. Salant(1), L. Cook(2)

(1) Georgia Institute of Technology

(2) Rodel, Inc.

ABSTRACT

Chemical-Mechanical Polishing (CMP) refers to a material removal process done by rubbing a work piece against a polishing pad under load in the presence of a chemically active, abrasive containing slurry. CMP is a commercially important process for planarizing silicon dioxide insulating films on partially fabricated integrated circuits produced on silicon wafers. Predictable planarization is an important step in fabricating denser and more reliable integrated circuits. CMP has been developed and used with empirically derived process procedures. The fundamental material removal mechanisms are not well understood. This paper describes experiments to help determine fundamental mechanics of CMP.

CMP works through a combination of chemical dissolution and mechanical action. The mechanical action of CMP involves tribology. The liquid slurry is trapped between the work piece (wafer) and tooling (polishing pad) forming a lubricating film. Contact mechanics, pad and chemical reactivity contribute to material removal. To understand this mechanical interaction experiments were done with commercial 100mm diameter <100> silicon wafers to measure differential wafer displacement, drag forces and slurry pressures in the range of typical industrial polishing conditions.

Negative differential wafer displacement measurements and sub-atmospheric slurry pressure measurements revealed that the wafer compressively deforms the pad beyond its measured deformation from static normal load during polishing as a result of fluid and mechanical interaction. This additional deformation was as great as 25 microns. The experiments revealed high friction coefficients (> 0.4) for slurries, indicating that asperity contact is substantial. Friction coefficients for water were greater (~ 0.8) indicating surface reactivity with the CMP slurries.

INTRODUCTION

Polishing employs an abrasive suspended in a fluid slurry to remove precise amounts of material from a work piece. The work piece is brought in close contact with the slurry abrasive by a normal load against a polishing pad. Relative motion of the pad and work piece remove material.

Chemical-Mechanical Polishing (CMP) refers to a process where the chemical activity of the slurry works in combination with mechanical action to effect material removal. CMP has many uses in electronic material processing, optics finishing and other applications. This research has focused on CMP used in the planarization of interlayer dielectric (SiO_2) on partially fabricated integrated circuits on single crystal silicon wafers. Dielectric planarization is

important in economically fabricating denser and more reliable integrated circuits. Although this application is rapidly evolving, a basic understanding of the process is still lacking. Improved understanding can lead to quicker product development as well as lower integrated circuit fabrication costs.

The object of this research is to understand the mechanical interactions of CMP at the conditions typical of dielectric planarization.

Past research has examined whether asperity contact or slurry fluid dynamics bring the abrasive in contact with the work piece to effect polishing. Using analysis, Cook [1] showed that abrasive particle impingement could effect material removal in glass polishing. Runnells [2] showed by analysis that CMP, as used in dielectric planarization, occurs with the work

piece separated from the polishing pad by an elastohydrodynamic slurry film thickness. These two papers suggest that fluid dynamic interactions could define the mechanical effect of CMP. Cook [3] showed that material removal rate by CMP follows a Preston relationship where removal is proportional to normal load and speed. See Preston [4]. Cook showed that the Preston relationship indicated that asperity contact was required for interlayer dielectric CMP. Nakamura [5] had measured slurry fluid film thickness from differential wafer displacement for silicon wafer polishing and found that it had a direct effect on removal rate and planarity. However, these results were for very different pad materials. Also, these tests were done with laboratory scale silicon wafer samples (20 mm in diameter). Levert [6] measured fluid film thickness from differential wafer displacement for several pad materials under various polishing conditions (load and speed). The measurements indicated that purely hydrodynamic separation of the work piece and pad can occur under small normal loads with commercial scale samples (100 mm diameter silicon wafers). This body of past research had indicated that mechanical and fluid dynamic interactions were both important in dielectric CMP, but this research did not define that interaction under typical industrial dielectric CMP conditions.

This paper reports on some basic parameters that are vital to an understanding of the mechanical and fluid interactions of CMP. This paper will give the results of differential wafer displacement, drag force and slurry fluid pressure measurements at the interface between the pad and work piece.

EXPERIMENTAL DETAILS

Experiments were performed on a Logitech PM-4 polishing machine with a 300 mm diameter horizontal table. Samples and measurement sensors were supported by a rigid overhead structure. Figure 1 gives a schematic diagram of the experimental apparatus. Polishing pads, supplied by Rodel Inc., were bonded to the turntable with pressure sensitive adhesive. The overhead structure maintained the wafer at a stationary position only allowing vertical travel

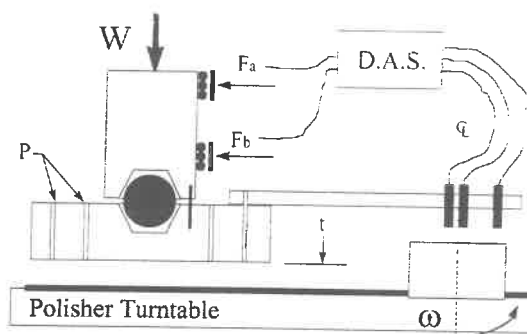


FIGURE 1 :
Schematic Diagram of Experimental Apparatus

and gimbaling of the sample mount. Silicon wafers were wax mounted to this rigid sample mount, which was designed to prevent flexure under test loading. Relative motion between the sample (silicon wafer) was accomplished by turntable rotation (ω). Turntable speed was electronically controlled. Normal load was applied by adding weights (W) to the overhead structure. This load was translated to the sample mount through a linear bearing which was restricted to vertical motion only. Slurry delivery was set at a uniform value with an automatic feeder.

Commercial 100 mm diameter silicon wafer samples were used for all tests. All polishing pads were comprised of a typical industrial perforated impermeable cast polyurethane material with a uniform thickness of approximately 1 mm. Deionized water and two CMP slurries were used. The first (base) slurry was colloiddally suspended silica in a high (~ 10.5) PH aqueous solution of NH_4OH . This slurry had a nominal 13% solids content and was typically used in CMP applications for polishing of SiO_2 dielectric layers on partially fabricated integrated circuits. The second slurry (20% solids) was similar except that it had a 20% solids content. The typical particle size for both slurries was 0.2 microns.

Measurements were taken for differential wafer displacement (t) with three capacitance air gap probes obtained from Capacitec. These probes were cantilevered from the sample mount to suspend the probes above a reference surface which was bonded to the rigid turntable. Three probes were used so that the entire wafer range of wafer motion and consequently differential wafer displacement could be detected. The probes have a sub-micron resolution.

Drag force measurements (F) were taken from the linear bearing mounts on the overhead structure using electronic load cells obtained from Entran. The combined readings from two load cells were used to remove the resultant moment. Both load cells were bolted into the structure with a known preload to maintain the stiffness of the overhead structure.

Measurements from the three capacitance probes and two load cells were automatically captured on separate channels with a PC based data acquisition system obtained from Data Translation Inc. This data was electronically downloaded to PC based data reduction software for plotting and analysis.

Differential wafer displacement was defined as the relative difference of sample position during CMP as compared to a static condition under the same normal load. This was done with sequential measurements as shown by the schematic diagram in Figure 2. The first measurement

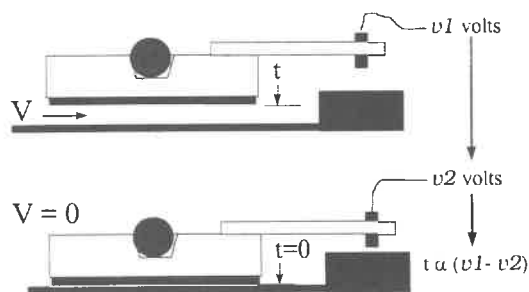


FIGURE 2:

Diagram of Differential Wafer Displacement Measurement Technique

was taken during the CMP experiment with the turntable in motion. A voltage $v1$ proportional to the air gap between the probe and the reference surface was acquired and stored by the data acquisition system. The turntable was stopped and the sample allowed to come to rest. The probe, which is rigidly attached to the sample, will also settle to its position of rest. A second reading $v2$ was then taken. The resultant differential wafer displacement was then proportional to the difference in probe voltages ($v1 - v2$) and several geometry constants.

Differential wafer displacement and friction measurements were taken at all combinations of 12 speed values, three normal load values and three fluids. Measurements were taken for the 12 speeds in random order for each of the nine combinations of normal load and fluids. Each of these nine combinations were also done in a random order. Differential wafer displacement measurements were taken in the following sequence. The zero wafer displacement ($v2$) was taken at six evenly spaced and randomly ordered locations which covered the swept area of the pad. Next the CMP wafer displacement ($v1$) was measured for four speed values in a short time span. Finally a repeat measure of zero wafer displacement ($v2$) was taken at six the evenly spaced locations on the pad.

Wafer displacement data was reduced by averaging the CMP values ($v1$) over a 30 second, 166 hz acquisition. Likewise the zero wafer displacement values ($v2$) were averaged over the six static locations. The effective zero value was then linearly interpolated from the zero values ($v2$) taken before and after acquiring CMP wafer displacement values ($v1$).

Conditioning was done after taking measurements for four speed values. Conditioning was done for approximately 1 minute for every 2 minutes of running time with a hand held diamond grit tool. Conditioning refers to the roughening of the pad by scratching the surface with a fixed multiple diamond grit tool. Pad wear and accumulation of slurry particles and wear debris tend to smooth the pad surface and change the polishing rate. Conditioning maintains a consistent pad surface texture to prevent changes in polishing rate.

Pressure measurements were done with either water column manometers or an electronic pressure calibrator. Pressure measurements were taken in random order and were allowed to stabilize before being recorded. Equipment limitations prevented conditioning which was completely consistent with that maintained during fluid film and drag force measurements.

RESULTS

The results of differential wafer displacement measurements are given in Figure 3.

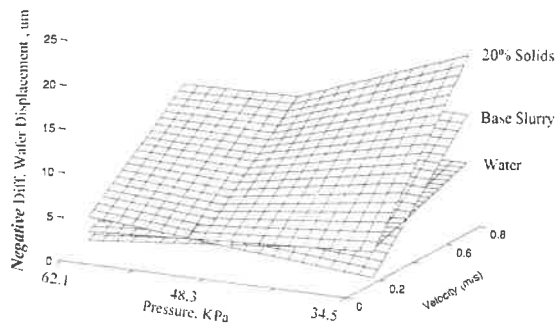


FIGURE 3:
Measured Differential Wafer Displacement.
Note all Values of Differential Wafer
Displacement Are Negative.

The three surfaces are a non-linear regression of 108 data points. The **negative** value of measured differential wafer displacement is plotted along the vertical axis while unit normal load and speed are plotted along the two horizontal axes. Speed was taken as the polishing machine tangential turntable speed at the wafer center. All measured differential wafer displacement values were negative for all tested combinations of normal load speed and fluid. Differential wafer displacement measurements became more negative as speed increased and the solids content and viscosity of the slurry increased. Differential wafer displacement measurements became more negative as the unit normal loading decreased.

A plot of the pressure of the fluid (water) trapped between the polishing pad and sample is shown in Figure 4. The slurry pressure, as a percentage of unit normal loading, is plotted on the vertical axis. The location of the pressure was measured the wafer is plotted on the horizontal axis where 0.0 is the inlet side and 100mm is the exit side across the center of the sample. The slurry pressure was measured to have both positive and negative (gauge) pressures depending on the location.

The average slurry pressure beneath the sample was 5% of unit normal load. This was obtained by integrating the curve. It must be noted that the results of the pressure curve can not be fully

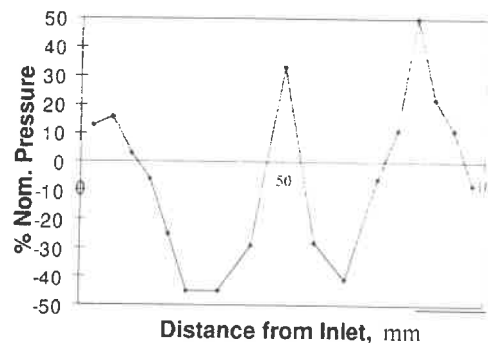


FIGURE 4:
Measurement of Slurry Pressure Between Wafer
and Polishing Pad for Water at 34.5 KPa Unit
Normal load and 0.45 m/s.

compared to the other results because consistent pad conditioning was not possible during the pressure measurements. Also, the pressure taps were rather large (1.0 mm in diameter) as compared to the actual scale of the pad asperities (~ 0.10 mm). The geometry of these asperities had a marked effect on fluid film measurements and it is likely that they may also affect pressure measurements as well.

Drag force is given as a curve of friction coefficient in Figure 5.

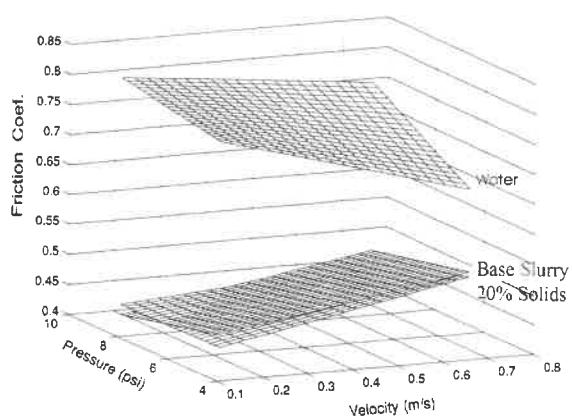


FIGURE 5:
Measurements of Friction Coefficient.

The three surfaces are a non-linear regression of 108 data points. Friction coefficient is plotted along the vertical axis while unit normal load and speed are plotted on the two horizontal axes. Both slurries have relatively high (~ 0.4) friction coefficients. The friction coefficient for water is almost double that for the slurries (~ 0.8). The friction coefficients for all three fluids did not vary greatly with normal load. Friction coefficient values for water declined slightly with increasing speed while the values for both slurries increased slightly with increasing speed.

DISCUSSION

The high coefficient of friction measurements (Figure 5) indicated that substantial asperity contact exists under all the conditions tested. Full hydrodynamic separation of the polishing pad and slurry would have given friction coefficients an order of magnitude less than the measured values. The greater coefficient of friction for water indicated that both slurries significantly modified the polishing pad and/or wafer surfaces to reduce friction under all conditions tested.

The experiments indicated that the differential wafer displacement (Figure 3) between the polishing pad and wafer was less during CMP than it will be for an equal statically applied normal load for all conditions tested. The substantial asperity contact, along with negative differential wafer displacement, indicated that pad asperity compressive deformation is greater during CMP than with a statically applied normal load. This additional compressive asperity deformation requires that an additional force be present which further pushes the wafer into the polishing pad. The sub-atmospheric slurry pressure measurements (Figure 4) revealed a source of this additional force.

The polishing pad geometry (surface topography and permeability) and polishing conditions, such as normal load and speed, can have a strong effect on CMP. The test conditions of normal load and speed, the perforated impermeable polyurethane polishing pad, the slurry and the conditioning level, used in these experiments, were all typical of CMP done for SiO_2 dielectric planarization for partially fabricated integrated

circuits. However, different polishing pad materials and polishing conditions can yield strikingly different results. Earlier research [6] indicated that large, positive, separating fluid film thickness can be present during polishing for an unperforated version of the same pad for normal loads at 10% of the current test values. This earlier research also indicated that increased polishing pad permeability can greatly affect the values of differential wafer displacement.

PROPOSED CMP MECHANISM MODEL

Polishing pad asperity contact is required for SiO_2 dielectric CMP. In some cases, asperity deformation can be greater during CMP than in the statically loaded case because of a combination of parameters (normal load and speed) and polishing pad types.

It is proposed that this greater deformation may be caused by a repetitive "suction cup" effect as shown in Figure 6.

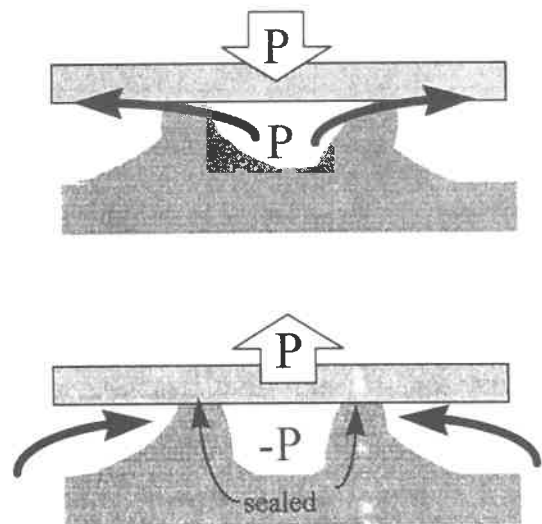


FIGURE 6:
Schematic Diagram of "Suction Cup" Effect.

The incompressible liquid slurry fills the pad perforations and pockets between pad/wafer asperity contacts expelling any gases. At the inlet region when the slurry is initially drawn beneath the wafer by pad rotation, pad/asperity compressive deformation can push fluid out of these pockets. Towards the center of the wafer, a relaxation from pad/wafer waviness or pad

relaxation from pad/wafer waviness or pad material time dependent modulus change can cause an expansion of these fluid pockets. The slurry pressure drops because the slurry is incompressible and cannot expand to re-fill the expanded volume of the pocket. Additionally, the rim of the pocket forms an effective seal which coupled with the impermeable pad prevents lateral slurry flow to restore fluid pressure. Without conditioning the pad surface becomes glazed. This limits asperity contact and promotes a greater hydrodynamic fluid film thickness. This larger film thickness allows for greater fluid migration limiting the chances of a negative (gauge) pressure build-up.

The negative differential wafer displacement measurements and negative (gauge) slurry pressure measurements were conclusive. However, the mechanism behind this "suction cup" effect is not well understood and this proposed model is not definitive.

CONCLUSION

Pad asperity contact is significant at conditions typically used in commercial CMP for dielectric (SiO₂) planarization on partially fabricated integrated circuits. This contact can be modified by polishing pad geometry and texture as well as changing polishing conditions (normal load and speed).

ACKNOWLEDGMENTS

The authors would like to thank Rodel Inc. for partial support of Joseph Levert during the summer of 1996 and for assistance in performing this research.

BIBLIOGRAPHY

- [1] L.M. Cook, *J. Non-Crystalline Solids*, v. 120, pp. 152-171, 1990
- [2] S. R. Runnels, *J. Electrochem. Soc.*, v. 141, No. 7, 1994
- [3] L.M. Cook, *Semiconductor International*, pp. 141-144, 1990
- [4] F. W. Preston, *J. Soc. of Glass Technology*, pp. 214 - 256, Vol. 11, 1927
- [5] T. Nakamura, *Bulletin of Japan Society of Precision Engineering*, v. 19, No. 2

- [6] J. Levert, *Transactions of International Tribology Conference - Yokohama* (Accepted for publication 12/95)

NOTE:

The identification of commercial materials (equipment & consumables) is given only for the sake of clarity. In no instance does such identification imply endorsement by Georgia Institute of Technology, or Rodel Inc. nor does it imply that the particular materials are necessarily the best available for the described purpose.

3D PROFILE GENERATION BY DOT-MATRIX METHOD USING MINIATURIZED ELECTRODE FEEDING DEVICES

Katsushi Furutani, Toshio Enami*, Naotake Mohri

Toyota Technological Institute

12-1, Hisakata 2-chome, Tempaku-ku, Nagoya 468 JAPAN

E-mail: furutani@toyota-ti.ac.jp

Abstract

This paper deals with a new prototyping method by electrical discharge machining (EDM). In this method, shapes are machined on a workpiece by scanning bundled wire electrodes which length is controlled independently. A prototype of a machining unit by this method has six miniaturized electrode feeding devices. The wire electrodes of 0.3 mm in diameter are arranged with the pitches of 0.76 mm. By compensating the wear of the electrode during EDM, various shapes with the accuracy of micrometers order can be obtained without machining a tool electrode.

1. Introduction

Recently, flexible machine tools with high accuracy are required to cope with various kinds of products which amount is small. Electrical discharge machining (EDM)^{1,2)}, stereolithography³⁾ and cutting⁴⁾ have been studied to make a fine structure of millimeters order in a short time. The current resolution of the stereolithography is sub-millimeters order and only resins can be used for material. Since shapes can be directly machined on metal by EDM without burrs, EDM is one of the available ways to machine microstructure.

EDM by scanning an electrode with a primitive shape has been proposed to generate three-dimensional shape without the shaping process of a tool electrode⁵⁻⁷⁾. As these processes are similar to machining with a ball end mill on a machining center, they should be flexible. As the peak current increases in EDM, the removal rate becomes larger. However, the maximum peak current is restricted in these systems to avoid continuous arc and rough surfaces. Therefore, EDM with multiple electrodes is more available to reduce the total machining time⁸⁾.

It has been proposed to use bundled rods to form three-dimensional shapes^{9,10)}. In these systems, positioning the rods takes for a long time because those units have only one feeding device. It is sometimes

difficult to position the rods because of friction. Since the bundled rods have the same area as a workpiece, the number of the rods becomes enormous in case of a large workpiece.

The number of electrodes can be decreased by scanning several electrodes on the workpiece and the positioning time can be reduced by feeding each electrode independently. Hereafter, this method is called a dot-matrix method.

In this paper, a dot-matrix EDM is proposed and a prototype of the machining unit is described. Then its machining performance is investigated.

2. Concept of dot-matrix method

Fig. 1 shows a concept of the machining unit of the dot-matrix method. The machining unit consists of miniaturized electrode feeding devices of a direct drive type, an electrode guide and wire electrodes. The lower ends of the electrodes are bundled at the electrode guide with small gaps. The machining unit is mounted on a main axis of an electrical discharge machine. A workpiece is mounted on an xy NC table for scanning motion. The machining process by the dot-matrix method is similar to a printing process with a dot-impact printer. This dot-matrix method can be applied not only to EDM but also to electrochemical machining and

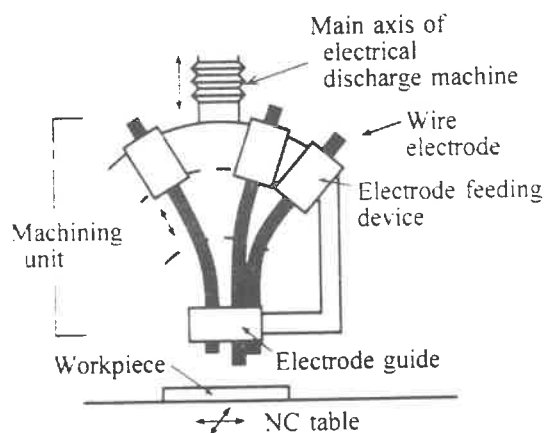


Fig. 1 Concept of machining unit by dot-matrix method

* Present address: Research and Development Center, Unitika Ltd. Uji, Kyoto 611 JAPAN

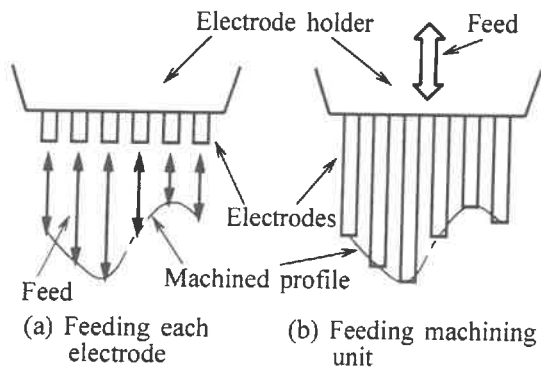


Fig. 2 Two types of feeding

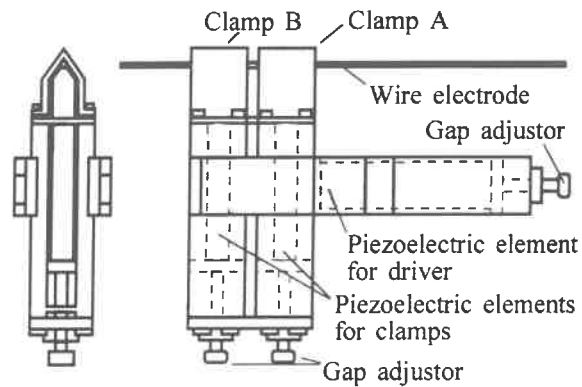


Fig. 4 Structure of Inchworm mechanism

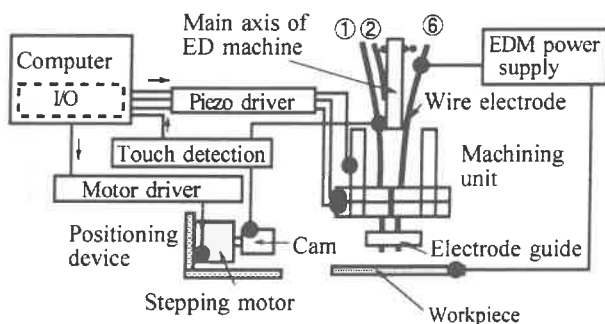


Fig. 3 Configuration of dot-matrix EDM system

forming.

Fig. 2 shows two types of feeding method:

- (1) feed all electrodes as one electrode
- (2) feed each electrode independently

The features of the dot-matrix method are as follows:

- (1) The length of the wire electrodes can be controlled independently with the electrode feeding devices.
- (2) The wear of the electrodes can be compensated by feeding them with the direct drive devices.
- (3) The scanning path of the machining unit is simple.
- (4) The projected area of the electrodes is much smaller than that of a workpiece.
- (5) The discharges can be forced to disperse since each electrode is isolated from the others. Consequently, EDM without cracks can be expected⁽¹⁾.

3. System for dot-matrix EDM

Fig. 3 shows a system configuration of the dot-matrix EDM. It consists of the machining unit, a positioning device, a personal computer and an electrical discharge machine. The machining unit is mounted on the electrical discharge machine. The machining unit and the positioning device are controlled with the personal

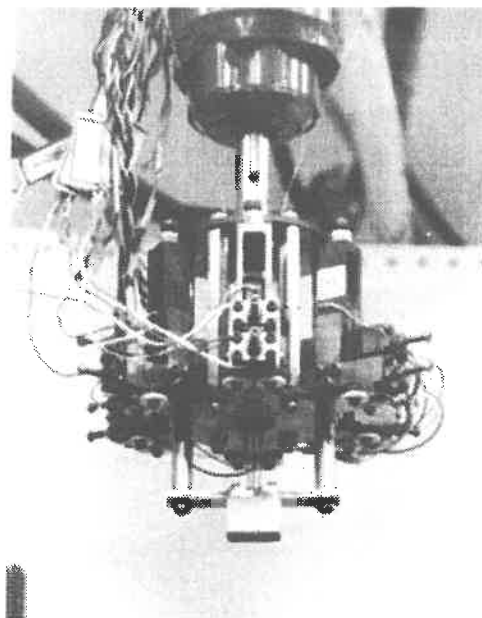


Fig. 5 Appearance of machining unit

computer (80386, 16 MHz).

To arrange the electrode feeding devices closely, the authors should make them miniaturized and thin. Miniaturized electrode feeding devices of a screw type⁽¹⁾, the Impact Drive Mechanism types⁽²⁾⁽³⁾, the Inchworm type⁽⁴⁾ and an elliptical movement type⁽⁵⁾ have been developed for EDM. After the comparison between these devices⁽⁶⁾, the Inchworm type is adopted in a prototype. The wear of the electrodes can be compensated with the Inchworm type because it feeds the electrodes directly.

Fig. 4 shows an electrode feeding device with the Inchworm structure. It measures $15 \times 44 \times 53$ mm. Multilayered piezos are used for clamps and a driver of the Inchworm. They deform $16 \mu\text{m}$ and $8 \mu\text{m}$ at 150 V, respectively. To adjust the clearance between an electrode and the clamps and the pre-load of the driver,

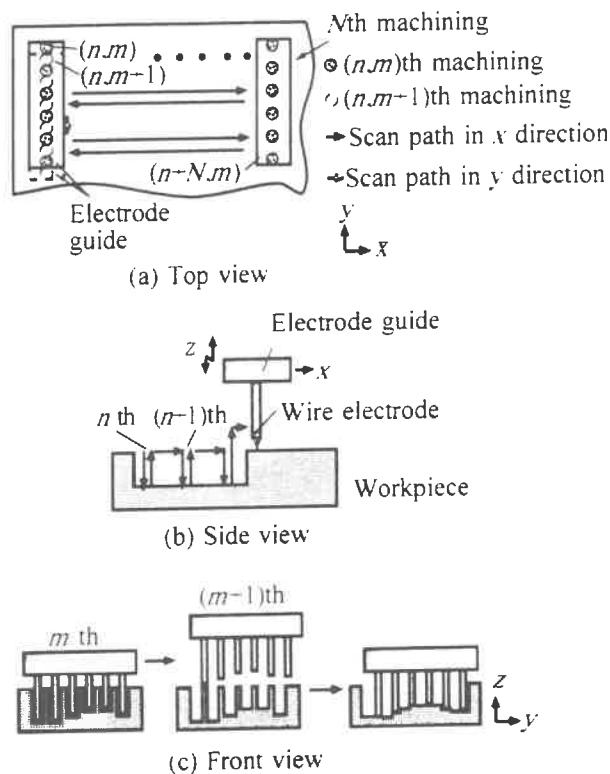


Fig. 6 Machining process by dot-matrix EDM

the positions of the piezos are arranged with screws. A tungsten rod of 0.3 mm in diameter is used for the electrodes. Steps of the electrodes are 1 μm under the conditions of 100 V as the applied voltage and the driving frequency up to 1 kHz.

Fig. 5 shows an appearance of the machining unit. It measures $55 \times 55 \times 80$ mm. Six electrode feeding devices are arranged on a circle. The electrodes are lined with the pitches of 0.76 mm at the electrode guide. As each electrode is isolated from the others, discharge can be forced to disperse.

Each electrode is positioned by detecting electrical contact to an elliptical metal cam which is driven with a step motor. The range of positioning is 0.5 mm with the resolution of 10 μm and the repeatable accuracy of 1 μm .

4. Three-dimensional profile generation

The machining process by the dot-matrix EDM is described and some examples of machined surfaces are introduced in this chapter.

4.1 Machining process

Fig. 6 shows the machining process by the dot-matrix EDM when the electrodes are fed as one electrode. The machining unit is scanned in the x and y

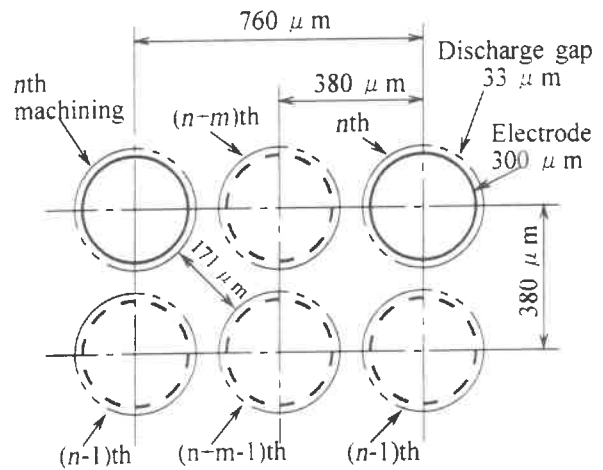


Fig. 7 Geometrical arrangement of electrodes

directions and machines in the z direction. The machining process by the dot-matrix EDM is as follows:

- (1) Adjust the length of the electrodes to compensate the wear of them.
- (2) Position the machining unit in the xy plane.
- (3) Machine shape by feeding the machining unit as an electrode.
- (4) When the machining is completed as shown in Fig. 6 (b), return to (1).
- (5) If the machining unit reaches the boundary in the x direction of the machining area, feed the machining unit in the y direction.

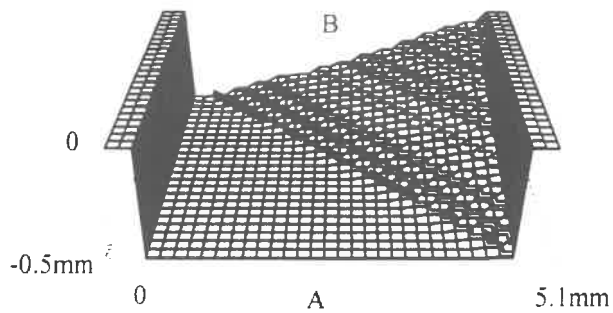
4.2 Machining condition

Shapes are machined under the low wear conditions as shown in Table 1 to reduce the error based on the wear of the electrodes. The electrodes are fed with the steps of 1 μm at the frequency of 200 Hz. A doped silicon plate with the resistivity of $7.52 \times 10^{-3} \Omega\text{m}$ is used for a workpiece.

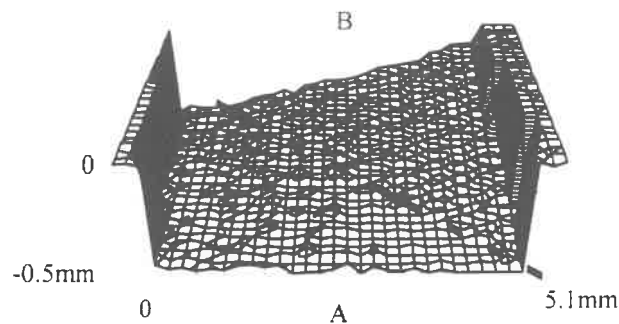
Since the gaps between the electrodes are larger than their diameter as shown in Fig. 7, it is required to make the pitch finer or to add a planetary motion in the xy

Table 1 Machining conditions

Gap voltage	150 V
Peak current	4 A
Pulse duration	16 μs
Duty factor	10 %
Radius of planetary motion	100 μm
Working fluid	EDF-K (Mitsubishi oil)
Polarity of electrode	-
Workpiece	Si ($7.52 \times 10^{-3} \Omega\text{m}$)



(a) Designed shape



(b) Result of Machining

Fig. 8 Example of machined plane and slope

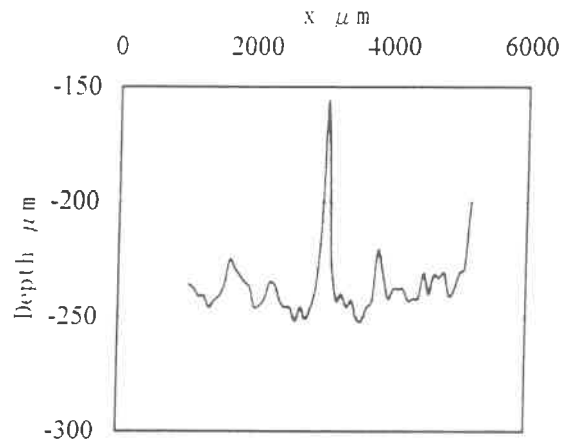
plane to machine without residual regions. In this case, the planetary motion is added to the feeding motion in the z direction. The required radius of the planetary motion is $86 \mu\text{m}$ because the gap is $33 \mu\text{m}$ under the conditions as shown in Table 1. Therefore, the pitches of scanning and the radius of the planetary motion are determined as $380 \mu\text{m}$ and $100 \mu\text{m}$, respectively. The electrodes are electrically connected in parallel.

4.3 Results of machining

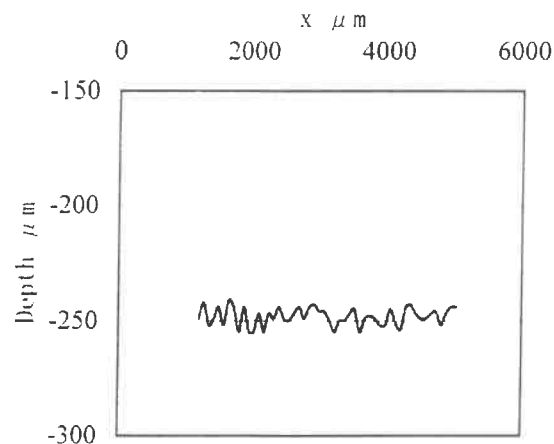
Planes and slopes are machined to investigate the performance of the dot-matrix EDM.

Fig. 8 shows a result of machined plane and slope. The machined area is $4.2 \times 4.2 \text{ mm}$ and the maximum depth is 0.5 mm . The result of machining has the same shape as the designed one. The height of the waviness is $20 \mu\text{m}$. The total machining time is 120 minutes including the positioning time of the machining unit. The ratios of the net time of machining time and the positioning time of the electrodes to the total one are 15% and 10%, respectively. The other time is spent in traversing between the positioning device and the workpiece. The total machining time can be reduced by automating.

Fig. 9 shows profiles of machined plane. The same



(a) After the first machining
(Average depth= $-240 \mu\text{m}$, $R_{\text{max}}=31 \mu\text{m}$)



(b) After the second machining
(Average depth= $-248 \mu\text{m}$, $R_{\text{max}}=14 \mu\text{m}$)

Fig. 9 Profiles of machined plane
(Designed depth= $250 \mu\text{m}$)

area is machined after compensating the wear of the electrodes. The machined area is $4.2 \times 3.8 \text{ mm}$ and the designed depth is 0.25 mm . Both total machining time of the first and the second machining are 30 minutes. After the first machining, the depth error and the height of the waviness except for a protrusion are $10 \mu\text{m}$ and $31 \mu\text{m}$, respectively. After the second machining, these errors are decreased to $2 \mu\text{m}$ and $14 \mu\text{m}$, respectively. The protrusion is eventually removed.

The wear of the edge portion at the end of the electrodes causes the waviness on the machined surfaces. As the result of machining, the machining error can be reduced by repeating machining on the same area with the wear compensation of the electrodes.

5. Conclusions

The dot-matrix EDM with scanning motion of the bundled electrodes is proposed in this paper. The conclusions are summarized as follows:

- (1) A prototype of the machining unit by the dot-matrix EDM with six electrode feeding devices of the Inchworm type is manufactured in trial.
- (2) Three-dimensional shapes can be machined by controlling the electrode feeding with the planetary motion.
- (3) The machining accuracy can be improved to micrometers order by repeating machining and compensating the wear of the electrodes.

The machining performance in case of feeding each electrode independently and the machining without cracks by forced discharge dispersion¹¹⁾ will be investigated.

Acknowledgement

The authors wish to thank to Prof. T. Higuchi of the University of Tokyo, and Mr. H. Misawa, Mr. S. Ishikura, Mr. N. Andoh and the staffs of the workshop of Toyota Technological Institute for helpful cooperation. This study is supported financially by OSG Fund, Amada Foundation for machine tools, Research Foundation of Electrotechnology of Chubu, and Grant-in-Aid for Scientific Research (B) (2) (06452196) and (A) (1) (06555041) of the Ministry of Education, Science, Sports and Culture of Japan.

References

- 1) Masaki, T., Kawata, K., Masuzawa, T. Micro Electro-Discharge Machining and Its Applications. *Proc. IEEE Micro Electro Mech. Sys. '90, Napa Valley, CA*, 1990, 21-26
- 2) Masuzawa, T., Fujino, M., Kobayashi, K., Suzuki, T. Wire Electro-Discharge Grinding for Micro-Machining. *Ann. CIRP*, 1985, 34(1), 431-434
- 3) Kruth, J. P. Material Increment Manufacturing by Rapid Prototyping Techniques. *Ann. CIRP*, 1991, 40(2), 603-614
- 4) Takeuchi, Y., Sawada, K., Sata, T. Computer Aided Ultra-Precision Micro-Machining of Metallic Materials. *Proc. IEEE Int. Conf. Robotics and Automation, Nagoya, Japan*, 1995, 67-72
- 5) Kaneko, T., Tsuchiya, M., Fukushima, T. Improvement of 3D NC Contouring EDM Using Cylindrical Electrode. *Proc. Int. Sympo. Electro Machining 9, Nagoya, Japan*, 1989, 49-52
- 6) Suzuki, K., Uematsu, T., Asano, S., Yanase, T., Ohkawa, K. An Attempt of Electrodischarge Machining on a Machining Center. *Proc. Int. Sympo. Electro Machining 9, Nagoya, Japan*, 1989, 38-41
- 7) Yuzawa, T., Magara, T., Goto, A., Imai, Y., Sato, T., Sendai, T., Kobayashi, K. Micro Contour EDM Using Thin Electrode (1st Report) - Evaluation of Fundamental Performance -. *Proc. Spring Annual Meeting of Jpn. Soc. Precision Eng., Tokyo, Japan*, 1995, 965-966 [in Japanese]
- 8) Furutani, K., Mohri, N., Higuchi, T., Saito, N. Development of Pocket-Size Electro-Discharge machine with Multiple Degrees of Freedom. *Proc. JSME Int. Conf. Advanced Mechatronics, Tokyo, Japan*, 1993, 561-566
- 9) Nakajima, N. Study on Dies and Electrodes with Bundled Wires. *J. Jpn. Soc. Mech. Eng.*, 1969, 72(603), 498-506 [in Japanese]
- 10) Finckenstein, E. v., Kleiner, M. Flexible Numerically Controlled Tool System for Hydro-Mechanical Deep Drawing. *Ann. CIRP*, 1991, 40(1), 311-314 [in German]
- 11) Yoshida, T., Mohri, N., Saito, N., Suzuki, M., Furutani, K. Monitoring Electrical Discharge Phenomena in Divided Electrode. *Proc. Spring Annual Meeting of Jpn. Soc. Precision Eng., Tokyo, Japan*, 1994, 905-906 [in Japanese]
- 12) Higuchi, T., Furutani, K., Yamagata, Y., Takeda, K. Development of Pocket-Size Electro-Discharge Machine. *Ann. CIRP*, 1991, 40(1), 203-206
- 13) Furutani, K., Mohri, N., Higuchi, T. Direct Drive Method of EDM Electrode Utilizing Impulsive Force. *Proc. Japan-France Congress of Mechatronics, Takamatsu, Japan*, 1994, 184-187
- 14) Furutani, K., Mohri, N., Higuchi, T., Saito, N. Direct Drive Method of EDM Electrode Utilizing Elliptically Moving Devices. *Proc. Machining Technology in Asian and Pacific Regions, Guangzhou, China*, 1993, 224-229
- 15) Morita, H., Furutani, K., Mohri, N. Electrical Discharge Device with Direct Drive Method for Thin Wire Electrode. *Proc. IEEE Int. Conf. Robotics and Automation, Nagoya, Japan*, 1995, 73-78
- 16) Furutani, K., Mohri, N., Higuchi, T. Miniaturized Electrode Feeding Devices Using Piezoelectric Elements. *Proc. Int. Sympo. Electro Machining 11, Lausanne, Switzerland*, 1995, 621-628

NANOMETRIC MEASUREMENT BY OPTICAL INTERFEROMETRY

Dr Michael John Downs
Division of Mechanical and Optical Metrology
National Physical Laboratory
Teddington Middlesex England

The development of lasers made optical interferometry the predominant technique for the high-precision measurement of displacement. These systems reversibly count the fringes produced by the interferometer, and in short-distance applications, they resolve to fractions of a fringe to achieve the required measurement precision. In most applications, these instruments are used in the free atmosphere, and to realize their potential accuracy - approaching 1 part in 10^8 - it is essential to correct the wavelength of the illuminating radiation for changes in the refractive index of air that otherwise might limit this performance by orders of magnitude.

Over short distances (less than $100\mu\text{m}$), the achievable accuracy becomes proportionally less sensitive to changes in the refractive index of the atmosphere, except the variations that cause anomalous path length differences between the two arms of the interferometer.

When subnanometric resolution is required, the lengths of the airpaths in the two arms of the interferometer must be minimized; a sufficiently accurate method for fringe subdivision must be used; and the fidelity of the signals generated from the interferometer must be maintained. Unfortunately, unwanted reflections and polarization effects cause fringe distortion and impose stringent limitations on the instrument when resolving to fractions of a fringe.

Michelson and Jamin type interferometers designed at the NPL to realise an accurate nanometric measurement capability will be described, as shown in figures 1 and 2 these include a cube-corner retroreflection system and a 'common-path' differential plane mirror system using 'hybrid' cube-corner plane mirror retroreflectors.

The limitations to the performance of these types of instruments when resolving to this precision will be discussed, together with the techniques developed to overcome the problems encountered in practice.

The results of the calibration of a high precision capacitance transducer shown in figure 3 will also be described.

Interferometer beamsplitter
characteristics

R_S = Reflectance (substrate/film
interface)

R_A = Reflectance (air/film
interface)

T = Transmittance

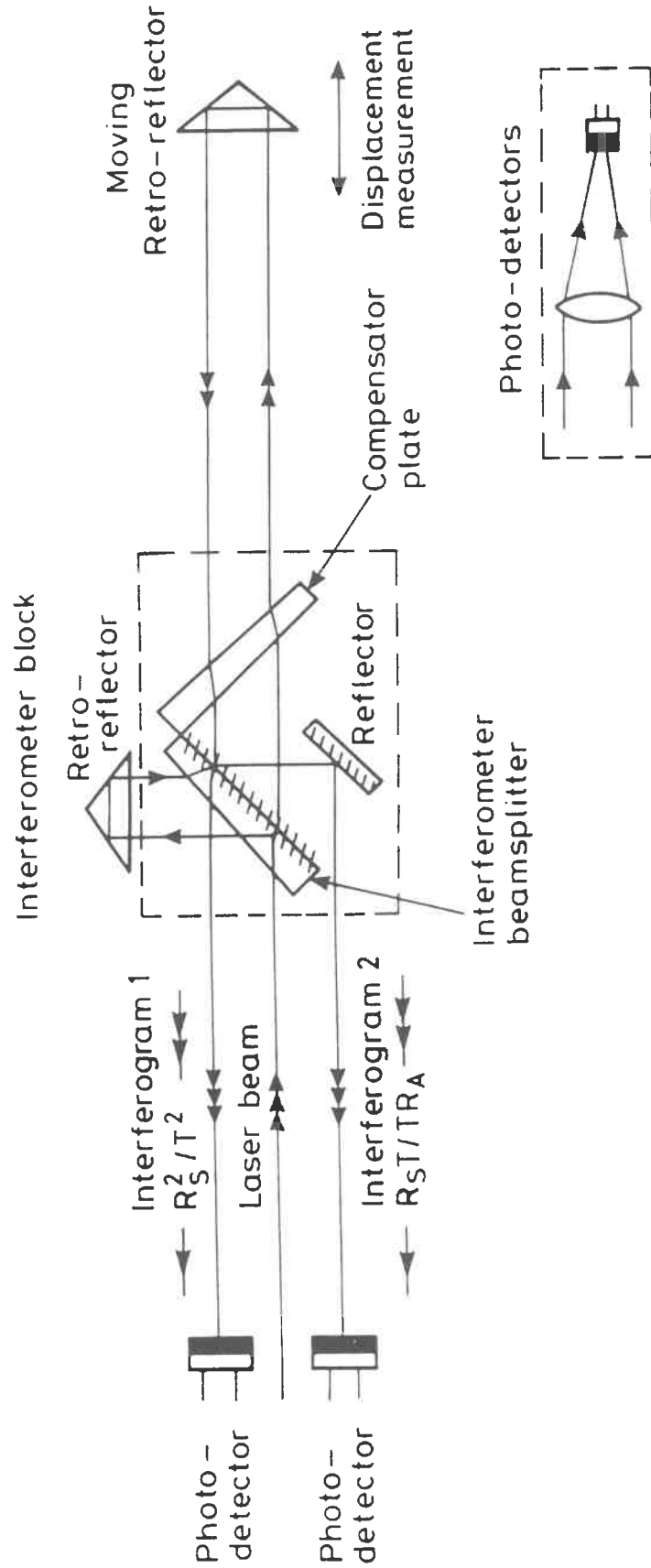
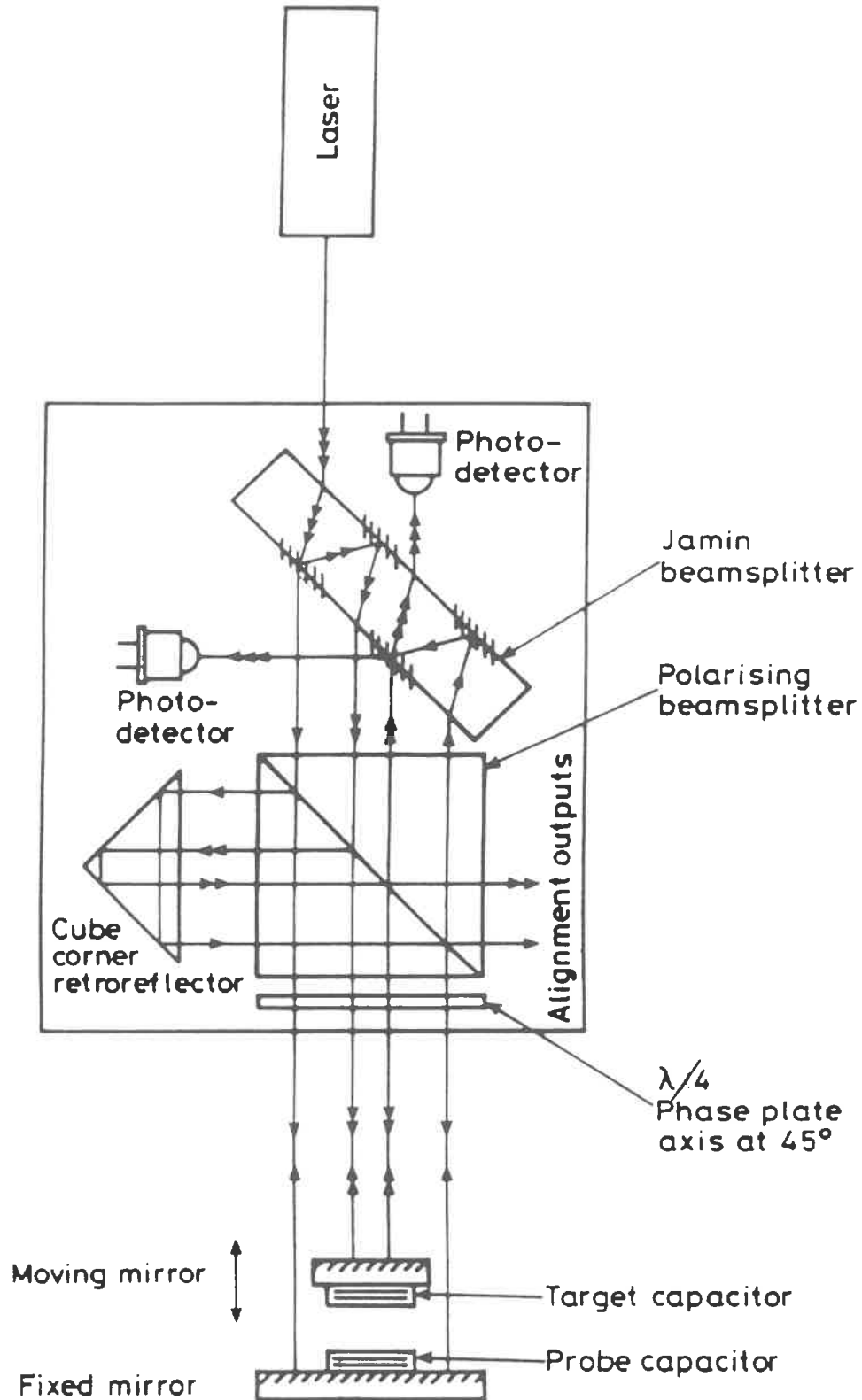


Figure 1

Cube Corner
Length measuring interferometer

Figure 2



Schematic of the NPL differential plane mirror interferometer calibrating a capacitance transducer.

Difference between Capacitance probe and Interferometer

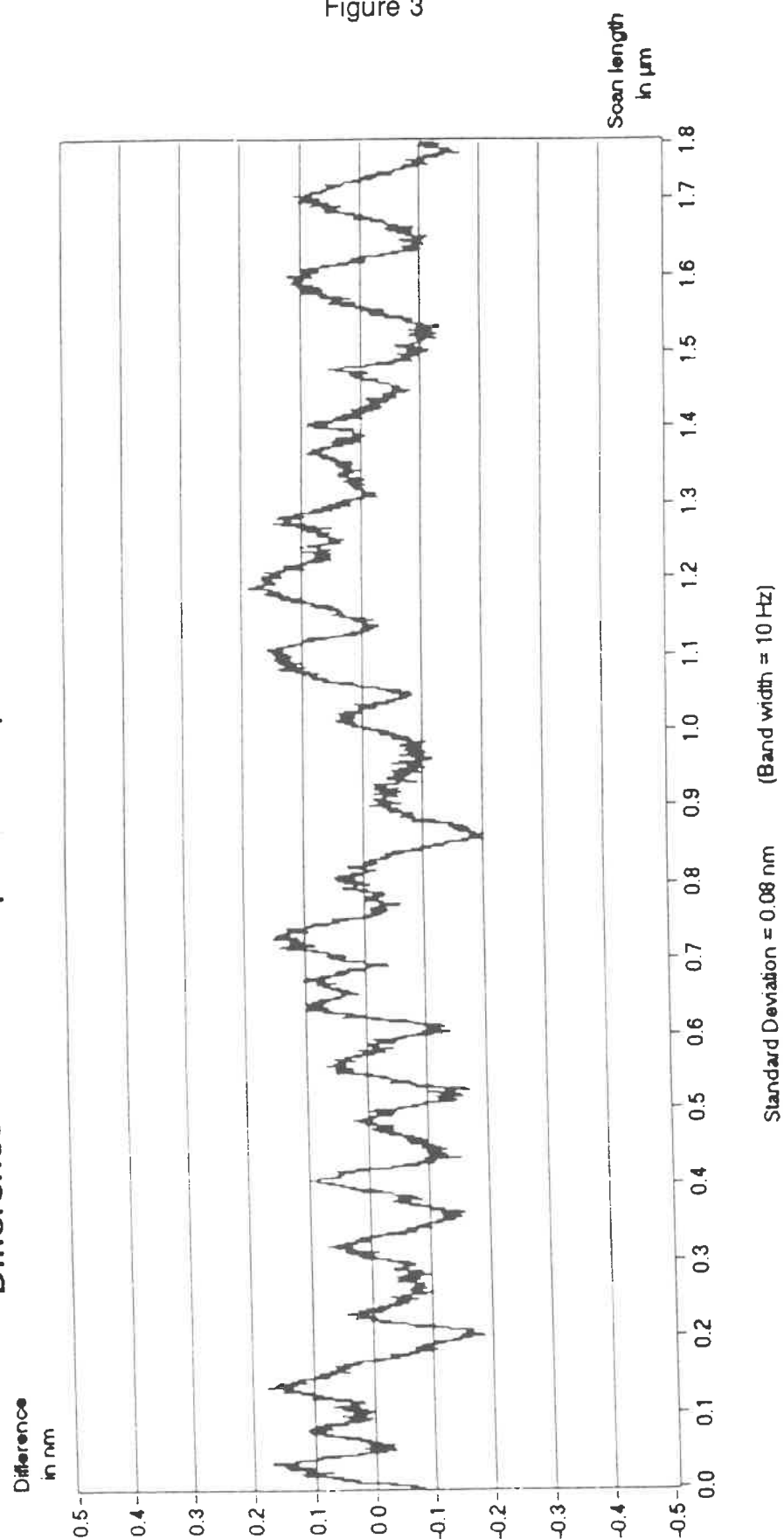


Figure 3

STUDY ON A NOVEL METHOD FOR REFRACTIVE INDEX COMPENSATION OF LASER MEASURING SYSTEM ON THE ULTRAPRECISION MACHINE TOOLS

F.Z.Fang, Y.Q.Zhang and G.X.Zhang

Tianjin University, Dept. of Precision Instrument Engineering

Tianjin, 300072 P.R. China

[Abstract] Research on the precision in-process measuring system with high measuring velocity is very important to the development of ultraprecision machine tools. One of the main factors balking the further enhancement of measuring accuracy of two-frequency laser is the changing of refractive index with environment, which produces the change of the wavelength, the measuring basis. The general method for solving the problem is using atmospheric parameter compensator. The method computes the refractive index using the atmospheric parameters, which is affected by the accuracy of atmospheric sensor, and the cost of the system is also high. This paper presents a new method. It uses a additional calibration optical path , which is put on the machine together with measuring optical path, to compensate the changing of refractive index. The influence of the refractive index on both path are similar, therefore it can be expected that the error caused by the changing of refractive index can be well compensated. The method can also compensate the optical distance change caused by the material temperature.

Keywords: Ultraprecision machine tool, Laser measuring, Refractive index

Introduction

Precision engineering has been playing a rapidly increasing and important role in manufacturing. Precision machining is one of the most successful development within precision engineering. An ultraprecision machine tool is the precondition of realizing ultraprecision machining, but without ultraprecision in-process measuring system, An ultraprecision machine tool could not machine ultraprecision components, especially precision optical element such as nonspherical surface and positive lens. Nowadays, the best project of realizing ultraprecision turning machine in-process measurement is the utilization of two-frequency laser measurement. Therefore study on enhancing the precision of in-process laser

measuring system becomes the key technology of development of ultraprecision machine tools.

Acousto-optical modulated (AOM) laser heterodyne measuring is not restricted by the small frequency difference.^{4,5} The principle of AOM is as follows. An ultrasonic light modulator driven by an ultrasonic signal source and power amplifier produces an ultrasonic wave, which passes through the medium (i.e. quartz, PbMoO₄, LiNbO₃, etc.) and causes refractive index n_s of the medium to change with sine rule. The distance between two sine waves is equal to the ultrasonic wave length λ_s , i.e. the medium is turned into the sine grating, with the rough and fine changing index. The grating space is λ_s . When the fine beam of the parallel laser is passed through this medium, the incident light is modulated by diffraction. The frequency of zero-order light is the same with the frequency of the incident light ν_0 , and the positive first order and negative first order frequency is individually $\nu_0 + \nu_s$ and $\nu_0 - \nu_s$. The positive second order and negative second order frequency is individually $\nu_0 + 2\nu_s$ and $\nu_0 - 2\nu_s$, and so on. Therefore, the incident light can be turned into a series of light beams through acousto-optical modulator with a certain of frequency difference, and it can be divided into separate light beams in space. Some proper measures are taken to restrict all of beams except zero-order and positive first order light, so that zero-order and positive first order light can pass through to the interferometric system. In this way, the frequency of ultrasonic signal source is adjusted so as to get the required heterodyne light beams of beat frequency.

Heterodyne laser measuring

A piezoelectric transducer (as shown in Fig.1) is one of the key parts of the acousto-optical devices, which turns the electric signal into the ultrasonic signal and interacts with the light. When the light passes through the diaphaneity light medium with the transmission of the ultrasonic wave inside, its diffraction characteristics is decided by the acousto-optic modulation. In a heterodyne laser interferometry, AOM acts as a grating, and diffraction fringes ejected can be used as coherent light. To the ultrasonic traveling wave produced by the electric signal $u(t) = U_0 \cos \Omega t$, the optic transmissivity function of the ultrasonic modulator is

$$T(x) = \exp[-i\alpha \cos(Kx - \Omega t)] \quad (1)$$

where, U_0 is the amplitude of the signal; $\alpha = kpcU_0l$ is the amplitude of the light phase modulation; $k = 2\pi/\lambda$ is the light wave number; p and c are the piezo-optical coupling

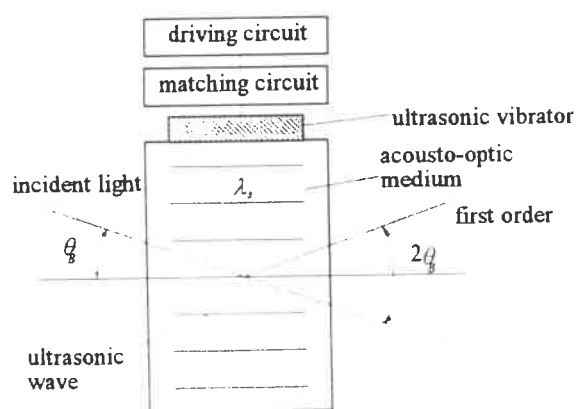


Figure 1 Principle and construction of acousto-optic devices

constant; $K = 2\pi/\lambda_s$ is the ultrasonic wave number; $\Omega = 2\pi\nu_s$ is the angular frequency of the ultrasonic wave; l is the depth of the ultrasonic field.

AOM forms the plane wave $E_1 = E_0 \exp(-i\omega t)$, where, E_0 is the amplitude; $\omega = 2\pi\nu$ is the light angular frequency. The diffraction light spectrum out of the modulator is directly proportional to the Fourier convolution of the transmissivity function and the AOM aperture function.

$$A(u) = \int_{-\infty}^{+\infty} E_1 T(x) \text{rect}\left(\frac{x}{B}\right) \exp(-i2\pi ux) dx$$

$$= E_0 B \sum_{-\infty}^{+\infty} J_m(\alpha) (-i)^m \exp(-i(\omega + m\Omega)t) \text{sinc}\left[(u - m\frac{1}{\lambda_s})B\right] \quad (2)$$

where, $J_m(\alpha)$ is the Bessel's first class function of m-grade; $u = \sin \alpha / \lambda$ is the space frequency; α is the angular displacement; $\text{rect}\left(\frac{x}{B}\right) = \begin{cases} 1, & |x| \leq B/2 \\ 0, & |x| > B/2 \end{cases}$ is the aperture function of

AOM; $\text{sinc}\left[(u - m\frac{1}{\lambda_s})B\right] = \frac{\sin \pi B(u - m/\lambda_s)}{\pi B(u - m/\lambda_s)}$; B is the filter pass band after the optical and electrical frequency being mixed.

From the equation (2), the optical field in AOM aperture is the sum of a group of coherent plane wave of the different direction angle $\alpha_m = m\lambda / \lambda_s$ and the different frequency $(\omega + m\Omega)$.

When the interaction is weaker ($\alpha < 0.5$), the light spectrum (7) includes the 0-order and ± 1 -order diffraction at least.

$$A(u) \approx E(0) + E(+1) + E(-1)$$

$$= E_0 J_0(\alpha) \exp(-i\omega t) - iE_0 J_1(\alpha) \exp[-i(\omega + \Omega)t] + iE_0 J_1(\alpha) \exp[-i(\omega - \Omega)t] \quad (3)$$

At this time $J_0(\alpha) \approx 1$, $J_1(\alpha) = b_1 kpcU_0 l$, where, b_1 is the constant, and the amplitude of ± 1 -order diffraction light is directly proportional to $E_0 U_0 \exp[-i(\omega \pm \Omega)]$. From the equation (3), the frequency modulation is the result of the acousto-optical signal frequency mixing in the modulator.

Compensation of refractive index

The error being caused by refractive index is the most significant sources of error in precision laser measurement. In making in-process measurements, it is necessary to make the correct compensation for the change in the wavelength of the laser light due to the ambient conditions of the air, and thermal expansion or contraction of the machine tool. The absolute accuracy of the laser interferometer is directly determined by how accurately the ambient

conditions are known; as a guide, it can be stated that an error of approximately 1 part per million will be incurred for each error of 2°F (1°C) in ambient temperature, 0.1-in. (2.5 mm) Hg in absolute pressure, and 30% in relative humidity. In general, the automatic compensator is intended to be used for in-process measurement purpose. In order to enhancing the compensation accuracy, many high precision air sensors are needed to get the real refractive index of the light path. Fig. 2 shows a novel method for refractive index compensation. It is simple and different from the compensation method of using the reference light path. 0-order and 1-order light beams from AOM come into measuring path and calibration path respectively. Both of them are arranged on the same light path at the same time, and interfere with each other. Reflector R2 is fixed. 0-order and 1-order light beams travel in both measuring path and calibration path, so the signal from the receiver 2 include the error being caused by the change of the refractive index. It is used to compensate the change of the refractive index of the measuring light path. Besides, the another most significant source of error is the influence of optical path difference caused by thermal changes. The error from the change of the material temperature is also compensated by this method. It is the effect of leadscrew to determine carriage position, this effect represents an expansion of approximately $6\text{ }\mu\text{m/m}$ for a 1°F rise in the leadscrew temperature. If the total carriage travel is 1 m, this effect represents a change in the leadscrew length of $6\text{ }\mu\text{m}$ per degree Fahrenheit change in temperature.

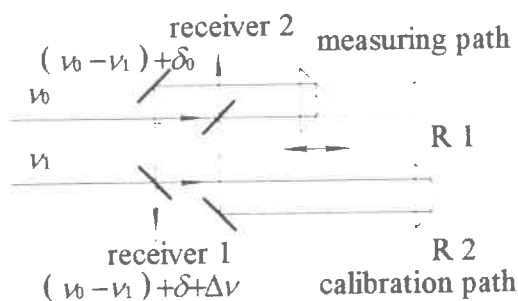


Figure 2 Principle of calibration path method

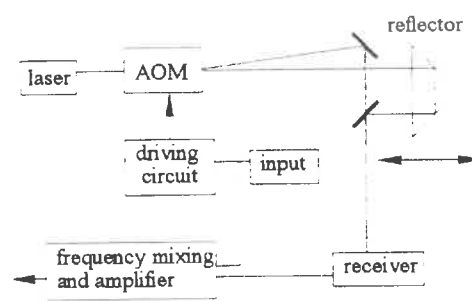


Figure 3 Principle of measuring system

Figure 3 is a principle sketch of realizing the AOM laser heterodyne measuring. The single frequency laser through the modulator produces 0-order and 1-order diffraction light, and the frequency difference between two beams of light is the ultrasonic frequency ν_s entering AOM. Doppler frequency shift $\Delta\nu$ produced by the movement of reflecting prism is received by the photoelectric receiver. The frequency shift $\Delta\nu$ carrying the displacement message is then taken out through a treatment circuit.

Experiment condition

The whole arrangement of machine tool is the T-base-lathe that axis of abscissa and axis of bank are separated (as shown in Fig.4), which provides a lot of convenience to the position and the measurement to reduce many error sources. For both the bed and the saddle are made of granite material, the position between each kind of lens and interferoscope is fixed

relatively. Special sealing system is used in the light path to reduce the influences of environmental turbulence in measuring path and stabilize the air parameters, and the calibration light path is also arranged at the same time, therefore the error of air refractive index is reduced obviously. The error in sliding way linearity and in perpendicularity between abscissa sliding way and bank one are given before as a system error to be compensated by the computer. The measuring moving mirror of bank slide carriage is fixed opposite cutter to make the mirror moving axis and the cutter moving axis as concordant as possible to eliminate Abbe error. Measuring moving mirror is fixed beside slide carriage of main spindle for the restriction of construction and then Abbe error is produced. This error may be eliminated by arranging symmetrically the light path in both sides of the main spindle. When the included angle between the spindle axis and the measuring axis is produced, Abbe error is produced from arranging one side light path. If two light paths are arranged symmetrically, Abbe errors of equal size and contrary direction are produced, and then eliminated by signal treatment.

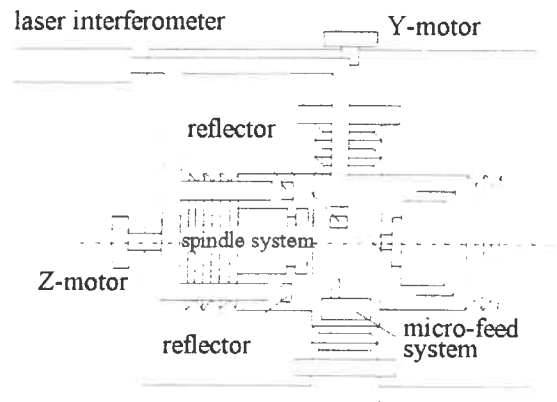


Figure 4 Arrange of the measuring system

Conclusion

An ultraprecision machine tool is a necessary equipment of realizing ultraprecision machining, and in-process laser measuring system is key to the development of an ultraprecision machine tool. In this paper, two-axis in-process measurement is realized on an ultraprecision turning machine, and a novel compensation method of refractive index is proposed. It can compensate not only the error from the change of refractive index effectively, but also the error from the change of material temperature. The measurement capability of the in-process measuring system is as follows, resolution 5 nm, range 1 m, measurement accuracy 0.15 ppm, max velocity 1 m/sec.

References

- 1 Fang, F.Z. and Yuan, Z.J. "Research on in-process measuring system of an ultraprecision turning machine," IEEE ICIT'94, 1994, 429-433
- 2 Fang, F.Z. and Yuan, Z.J. "Two-frequency laser measuring system used in ultraprecision turning machines and the enhancement in precision," The 5th Sino-Japan joint seminar on ultraprecision machining, 1994, 49-54
- 3 Fang, F.Z. and Yuan, Z.J. "Positioning and measuring technique on ultraprecision turning machines," China mechanical engineering, 1994, 5(2), 49-51
- 4 Gazalet, M.G., Ravez, M., Haine, F. and Bridoux, E. "Acousto-optic low-frequency shifter," Appl Opt, 1994, 33, 1, 293-1, 298
- 5 Sommargrem, G. "A new laser measurement system for precision metrology," Prec. Eng, 1987, 179-184

DETERMINATION OF STRAIGHT LINE PROFILES BY USING ANGULARLY DISPLACED CCD WITH PIXEL OUTPUT INTERPOLATION

Richard Jablonski

Shizuoka University, Department of Engineering, 5-1, Johoku 3-chome, Hamamatsu, Japan
Warsaw University of Technology, Institute of Metrology and Measuring Systems, Warsaw, Poland

Keywords: CCD array, shadow method, signal output interpolation, light intensity distribution, pixel geometry

1. Introduction

Image detectors based on two-dimensional CCD arrays nowadays enjoy widespread use in many applications ranging from scientific measurements and industrial research to on-line process monitoring. New high-sensitivity rectangular formats and increases in resolution elements continue to fuel this trend. For typical operations the manufacturers of imaging devices are offering a comprehensive line of CCD sensors and related modules but specific real-time measurements and analysis require special approach.

In measurements based on shadow method, where an object to be measured is illuminated (or scanned) by the laser beam - its shadow size is usually determined by the one-dimensional (1D) CCD array. Because mostly the cylindrical objects are measured, the shadow forms a band, bounded by two straight lines, which width corresponds to the diameter of measured object. The resolution is limited to the size of pixel and accuracy depends on the quality of array. When higher accuracy is required, the pixel output can be interpolated and then more detailed error analysis should be carried on, comprising not only the components resulting from array itself (like stability, pixel output linearity, geometry), but also from the inhomogeneity of laser beam, diffraction effects and position of an object in relation to the array. Besides, especially when small-pixel-size array is used - error caused by

dust particles must be taken into account.

The 1D CCD based technique is widely used in manufacturing processes - to control only the local diameter of an object moving with high speed, but when also the geometric form of an object has to be measured - the two-dimensional, small format, small pixel size CCD arrays are usually applied and or static measurement or measurement at low speed is carried on. Unfortunately small pixel size CCD's feature low uniformity level, and besides, even with smallest available pixel size CAD the attainable accuracy is not satisfactory. These problems can be solved by applying complicated calibration procedure based on modulation-transfer function measurement techniques [1,2], but its application to small format Cad is unjustified.

In the paper a simple method - significantly increasing the accuracy/resolution of CAD-based measurements is described. As a detector, a two-dimensional CAD, rotated through a certain angle with respect to shadow line (object profile), is used. Section 2 describes a technique for determining the optimum number of pixels and angular displacement. Section 3 provides the analysis of output signal interpolation. In Section 4 the experimental considerations are presented.

2. Method and theoretical analysis

Let us assume that the profile of an object represents straight line $y = x \tan \alpha + b$ on uniformly illuminated CAD, as shown in Fig. 1. The CAD is composed of m columns and n

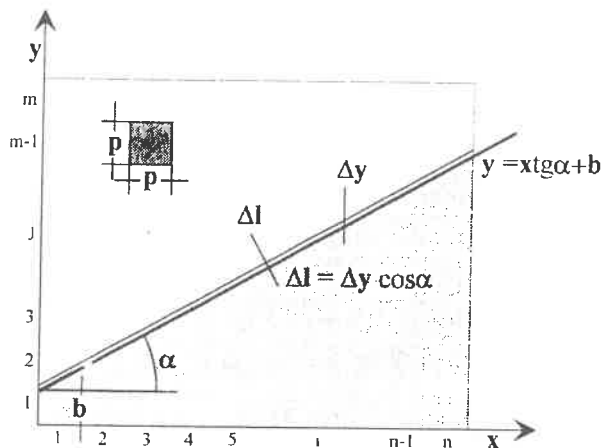


Fig. 1. Profile line and CAD geometry

rows of uniform square detectors, of a size $p \times p$ with negligible small spacing between them. The output signal for such situated object (shaded area) is:

i/j	1	2	3	...	j	...	m
1	1	0	0	...	0	...	0
2	1	1	0	...	0	...	0
3	1	1	0	...	0	...	0
4	1	1	1	...	0	...	0
5	1	1	1	...	0	...	0

(where 0 and 1 is the binary response of each pixel) and represents a staircase function of $p \times 2p$ steps. In general the function is formed by two superimposed periodical functions of different steps (steps of long period function are much bigger than these of short period ones). The size of a step and its periodicity depends on α . And now let us determine the smallest parallel displacement l of profile line causing 1 bit change of output signal. This dependance is an implicit function of 3 variables:

$$\Delta l = f(n, p, \alpha)$$

easy to solve numerically. Its graphic representation is shown in Fig. 2: for chosen number of pixels n an optimum angle α , for

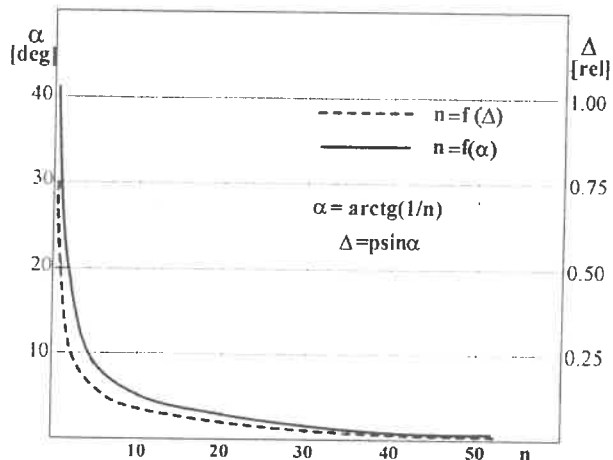


Fig. 2. Resolution vs. optimum angle α and number of pixels n .

which Δl reaches minimum, is calculated. E.g.: for $n=20$, an optimum angle is $\alpha=3\text{deg}$ and resolution $\Delta=0.08$ (Δ is a relative resolution $\Delta=\Delta l/p$). As it can be seen from Fig. 2, a high resolution is obtainable for small angles α , but according to Fig. 1, better results should be obtainable for α close to 45deg . This apparent contradiction is due to the periodicity of staircase function which in a big range favors small angles α . As it results from Fig. 3, the decrease of obtainable

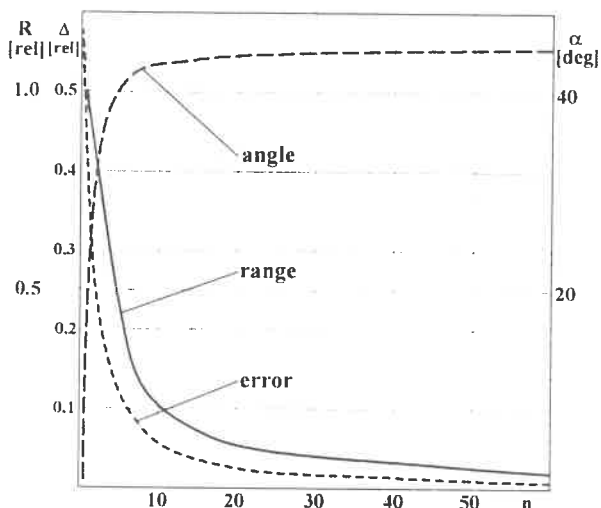


Fig. 3. Resolution and measuring range vs. optimum angle α and number of pixels n

resolution in a full range, in comparison to small range, reaches only $\sim 40\%$. Of course, for limited measuring range high resolution for

large α is obtainable; eg. for number of pixels $n=10$, optimum angle $\alpha=44\text{deg}$, expected resolution (error) $\Delta=0,05$, and measuring range $R=0.02$ (2% of full measuring range).

Fig.4 shows the resolution curves in full range

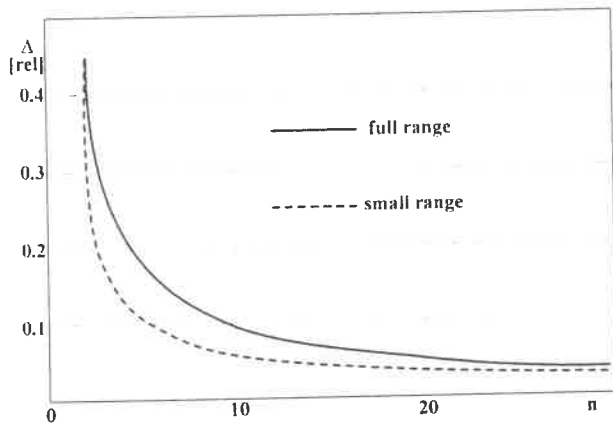


Fig. 4. High resolution (small range) and optimized resolution (full) range versus number of pixels

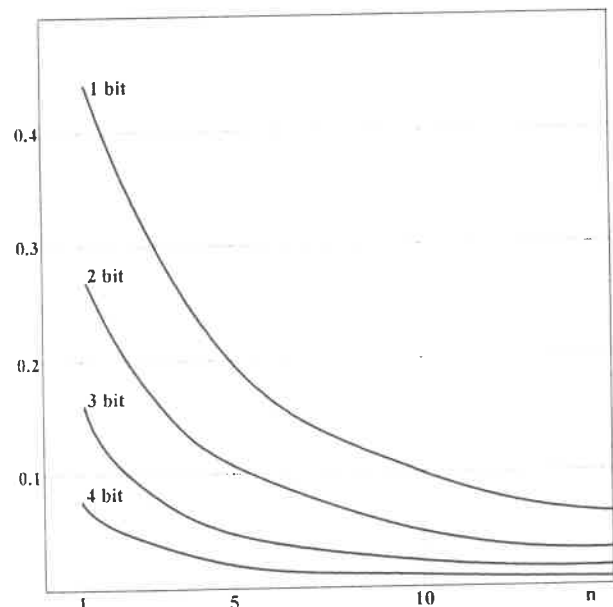


Fig.5. Resolution vs. number of pixels with pixel output interpolation

and in range within one period of long period function (- for big angles, such as on Fig.3).

3. Interpolation

In order to achieve higher resolution - the interpolation of pixel output signal was introduced. Fig.5 shows the obtainable resolution versus number of pixels for 1,2,3,4-bit interpolation.

Until now the discussion was limited to angle α and resolution Δ , and number of pixels n was taken optionally. Of course the bigger n the higher resolution is obtainable, but as it can be seen from above presented figures, for $n>8$ the gain in resolution rapidly decreases. On the other hand, larger n means longer measuring time due to the increased number of data for processing. For this reason $n=8$ was chosen as an optimum number of pixels.

Fig.6 shows an obtainable resolution for 1÷10 bit interpolation. One can easily separate the standard resolution range (4-bit interpolation) and high resolution range (8-bit interpolation). As interpolators the 12 bit A/D converters were used.

Equivalent results are obtained by introducing

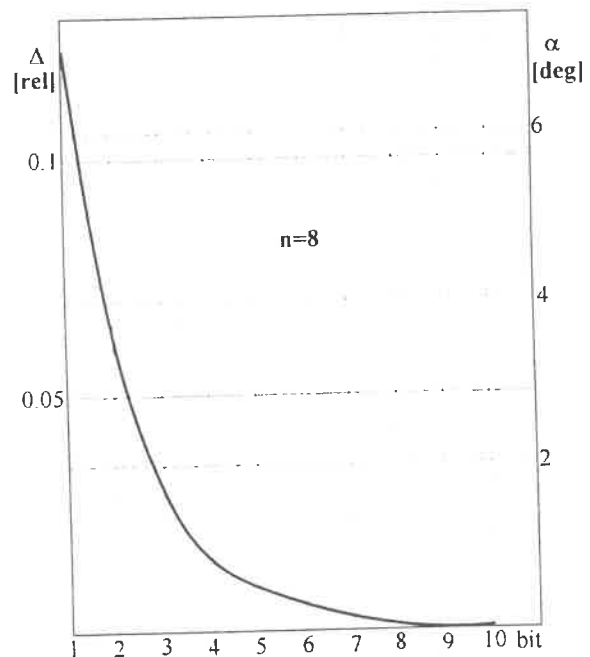


Fig.6. Resolution and angle α vs. number of bits

CAD of proportionally smaller pixel size. Eg. instead of 8 bit interpolation the $(1/256)p$ pixel size CAD.

Fig.7 shows the relationship between obtainable resolution and pixel size, for $\alpha=7.125\text{deg}$ and $\alpha=41.186\text{deg}$ - optimum

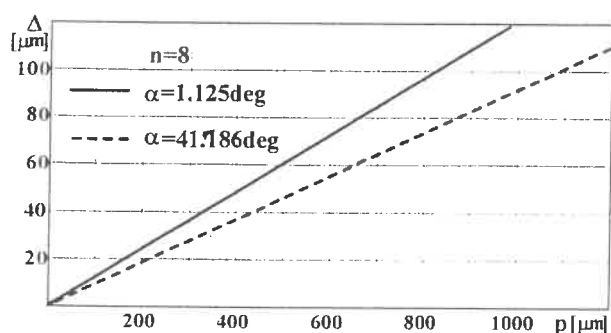


Fig.7. Resolution vs. pixel size (for $n=8$, $\alpha=1.125^\circ$ and $\alpha=41.186^\circ$)

angles for small and large angular range respectively.

4. Experimental considerations and results

An object was illuminated by the collimated He-Ne laser beam of known intensity distribution. The laser beam was expanded to 100mm diameter, its intensity distribution examined and 12x12mm area of most uniform intensity distribution chosen. For experiments Hamamatsu Photonics high-speed 16x16 element array of PIN silicon photodiode with sensitive area of 1.3x1.3mm was selected. The array was mounted on easy-to-use circuit board and each photodiode could be addressed individually.

The linearity (input/output characteristics) of each photodiode and also the illuminating light intensity distribution were measured, and the adequate corrections were introduced to the computer memory. Each pixel output was interpolated by 12 bit A/D converter and fed to PC for processing. The experiments were carried on 8x8 pixel area.

The experiments were focused on determining only one side of the shadow profile of cylindrical object.

The experimental set-up is shown on Fig.8.

The CAD was placed on the motorized y-stage fixed to the motorized rotating stage at the distance of 20mm behind the object. The best performance (minimal diffraction effects) was

obtained for Φ 3mm polished steel objects. The output was measured every 1 deg rotation angle within the 45 deg range and every 0.1 deg in proximity of optimum angle α . The displacements Δ were measured with

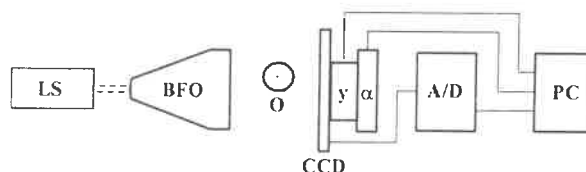


Fig.8. Experimental set-up. LS - laser. BFO - beam expander. O - object. y - motorized y-stage. α - motorized rotating stage.

accuracy 1 μm .

This results proved to be in good agreement with theoretical assumptions.

5. Conclusions

There are many CAD cameras designed for real-time measurements and image processing, and current PC technology allows continuous transfer of this digitized data directly to RAM. When higher resolution is required A/D converters are used but fast A/D rates are usually achieved with small-pixel, limited-dynamic range CCDs. Unfortunately this kind of CCDs characterize low geometrical accuracy and non-uniformity of sensitivity.

The described procedure enables to achieve accuracy 0.02 μm with 16 μm pixel size 8 column CAD and 8-bit interpolation. The procedure requires only ~7 deg rotation of CAD relative to profile line. It is sufficient to use 8-bit A/D converter for interpolation because further interpolation improves only slightly the resolution.

References

1. M. Marchywka, ... Applied Optics, Vol.31,1 Dec 1992, pp.7198-7213
2. M. Sasnsiper, ... Opt. Eng. 32, 1993, pp. 395-399.

THE DEVELOPMENT OF NON-CONTACT 3D PROFILE MEASUREMENT TECHNIQUE WITH PARALLEL PROJECTION BEAMS

Shuo-Jen Lee and Tien-Ron Fon
Department of Mechanical Engineering
Yuan-Ze Institute of Technology
Far East Road, NeiLi
Taiwan, R.O.C.
e-mail: mesjl@me.yzit.edu.tw

Abstract

In this paper, a non-contact 3D profile measurement technique was presented which is based on the projection of parallel beams. An arc light which was assumed to be point light source was positioned in the focal point of an parabolic mirror. After it was reflected by the parabolic mirror, uniform parallel beams were projected. These parallel beams, passing through LCD, produce 2^7 coded patterns. The purpose of these coded patterns is to ensure their relative positions. CCD images were then taken on these coded patterns in order to perform space coding on each of the measurement pixel points. It was followed by the calibration procedure with special designed model to obtain the internal and external parameters of the CCD such as focal length and ratio of pixel size. Using the triangulation theory, coordinate transformation between the measured object and the CCD will transform the 2D image plan into 3D coordinate profile. The parallelism of these projection beams eliminates the disadvantages of diffused projection light patterns, simplifies the mathematical algorithms, and reduces image processing error of the projected patterns. This can be more efficient and accurate technique for 3D profile measurement as compared with other non-contact structured light techniques.

Key words: 3D profile measurement, reverse engineering, CCD image processing, space coding, parallel projection beams, triangulation algorithm.

1. Introduction

3D Profile measurement of work piece by non-contact methods has been studied extensively because of its importance in the automated manufacturing processes and quality control of components.

Triangulation methods have been widely used in the non-contact profile measurement. The laser displacement probe(1) with positional signal detector which projects a laser spot on the object was comparable to the point by point contact type touch probe method. In contrast to the touch probe method(2), the measured position is the true coordinates of the object and did not requires touch probe radius compensation. However, its accuracy was greatly affected by the surface preparation of the object, the color of the object, the measurement range from the object and most of all from the geometrical parameters such as the slope and the curvature of the object.

The diode line laser which is a semiconductor laser with a pillar mirror which reflects the laser spot into a laser line(3,4). It improves efficiency and accuracy from the displacement laser probe. However, it have to incorporate CCD and image processing technique in order to digitize the laser's line signal. Various techniques were developed to process the laser's line signal and to calibrate the systems (5).

The structured light methods which incorporated binary patterns and CCD can process the area profilometry (6). Many projection methods and calibration algorithms were also studied in order to further improve the accuracy and efficiency from the line laser methods. However, for most of the line laser methods and the structured light methods, the size and shape of the dot laser beam or the length and width of a line laser beam of the projected slits were functions of the measurement range. They existed a fan pattern characteristics. The length and width of projected lines and patterns enlarge as they were further away from the light source. The accuracy and uncertainty of measurement increases with the measurement distance. A lots of efforts and caution have been conducted in the calibration process and mathematical manipulations are

required to assure the measurement accuracy.

In order to overcome these problems, we propose a new projection method and its calibration algorithm which theoretically should have infinite measurement range and at the same time maintain system accuracy. By incorporating space coding technique and triangulation algorithm, each measurement pixel point can be calibrated for its 3D coordinates on the 2D image. Experiments on the measurement of a semi-spherical model were conducted to verify the measurement procedures and the overall accuracy of the proposed method.

2. Principle of Measurement and Theoretical Considerations

The overall measurement system consists of (1) CCD camera; (2) projection system of parallel beams; (3) Liquid Crystal Display (LCD); (4) image processor; (5) personal computer; and (6) optical table. The system block diagram was illustrated on Figure 1.

The system components, their relative positions and coordinate relationships, as shown on Figure 2, were explained as follows. The local coordinates and their relative position of each system component can be analyzed by system setup parameters and coordinate transformation to compute the 3D absolute world coordinate of the measurement object from the 2D CCD image coordinate(7). The coordinate transformation from any one position to the projection view coordinate of the image plane is a many-to-one mapping. That is, in order to compute the true 3D world coordinate of a measurement point, one needs more than the information from the image plane coordinate to be able to transform the 2D coordinate to the 3D coordinate.

Coordinate Transformation

In this study, the LCD (Liquid Crystal Display) in the world coordinate projects a series of pre-determined Gray code patterns of which the x coordinates were known in advance. This is the necessary condition in this study for the 3D coordinate transformation. The geometry employed to derive the relationship between space code and CCD focal plane was illustrated in Figure 3. For the CCD image plane, the space relationship between the object and the image of the object on the CCD image plane was determined by the connecting line and the perspective plane projected by the Gray code. The connecting line from each pixel passing through the focal point of the CCD will have a unique intersection point with the plane of the Gray code. Because the connecting line passed through the focal point, the focal length of the CCD has to be determined accurately. This will ensure the correctness of the mathematical derivation and coordinate transformation between the 2D image plane and 3D object.

Calibration of CCD

The transformation parameters can be obtained by using a calibration plate with pins of known coordinates. In this study, a calibration plate with sixty-four known pin coordinates shown in Figure 5 was designed for this purpose. The calibration plate was positioned at the same measurement reference plane. CCD images of the calibration plate were taken with identical conditions and setup parameters such as focal length and fixed allocation between CCD and the object. The CCD images were processed for each of the 3D world coordinates of the calibration pins and their positions in the 2D CCD image plane. These data were used to compute parameters of the transformation matrix. The calibration procedures were as follows:

- (1) position the calibration plate in the measurement location;
- (2) open CCD and adjust focal length and focal lens;
- (3) take pictures of the calibration plate;
- (4) find the image centers of each of the calibration pins;
- (5) record center coordinates of the pins in the image coordinate;
- (6) compute parameters of the transformation matrix.

Generation of parallel projection beams

The pin-hole theory was employed to generate parallel projection beams. A arc light source was employed to emulate uniform point light source. A pin-hole positioned at the focal point of an off-centered parabolic mirror will project parallel beams. The generation of parallel projection beams was one of the key technologies of the proposed method. A series of procedures were also developed to calibrate the pin-hole position, the parallelism of the projection beams and the verticality of the CCD with the projection plate.

3. Measurement Procedures

A semi-spherical model with a diameter of 50.0 mm was designed and machined for this measurement test. The semi-spherical model despite of its simplicity has many desirable features. First, it is easy to machine. Secondly, the true coordinate of the whole model were known after the center of the ball was located. Third, it has varying slope, curvature and sharp edge w.r.t. the viewing CCD. Block diagram of the 3D profile measurement was shown in Figure 5. Figure 6 showed the original images and the threshold images of seven iterations of Gray code space coding. From the seven iterations of space coding, one hundred and twenty eight lines(2^7) were coded to cover the area of measurement.

4. Experimental Results

Figure 7 showed the picture of the measurement model and the line contours of the measurement image. Because there are one hundred and twenty eight lines to cover the area of 100.0(mm) X 100.0(mm). Therefore, the pitch between each line is about 0.8 mm. Taking the line that cross through the center, the average is about 0.8mm. This is due to the fact that the resolution of the CCD image is 0.8mm. Figure 8 indicated that the error is more significant along the sharp edge where the slope is close to ninety degrees. This is due to the fact the light intensity has the most drastic variation around that region. However, by adjusting the focal lens, the measurable slope is around 85 degrees. It is much better than most of the non-contact measurement techniques.

5. Conclusions and Discussion

The proposed method incorporated the simplicity of laser and the efficiency of area measurement of structured light. The generation and projection of parallel beams was developed for 3D non-contact profile measurement. Because the non-divergence nature of the parallel beams, the mathematical manipulation and coordinate transformation were much simplified. The accuracy of the measurement, in theory, is independent of the measurement distance. This is a significant advantage over other structured light method.

Because of parallel beams, the position of the LCD was very flexible. However, the penetration rate of the LCD employed in this study was lower than expected. This results in a dim projection intensity and may cause disconnected coded patterns of the Gray code when the edge of the object is closed to ninety degrees. This can be remedied by enlarging the focal lens. The results showed that the average error was 0.4 mm at the distance of 500 mm and forty three degrees with respect to the CCD. The resolution of the LCD also restricted the maximum number of projection lines to be one hundred and twenty eight. With the aid of another CCD, the measurement angle can be greatly improved. The other direction for improvement is to develop sub-pixel technique for the

References

1. F.H. Goh, et. al., "The applicability of a laser triangulation probe to non-contacting inspection," *Int. J. Prod. Res.*, Vol. 24, No. 6, 1986
2. M. Halioua and H. Liu, "Optical sensing techniques for 3D machine vision," *SPIE*, Vol. 665, 1986
3. R.C. Gonzalez and R.E. Woods, "Digital Image Processing," Addison Wesley, 1993
4. Wen-Chih Tai, Ming Chang and Shang-Hsiun Chen, "Non-contact five-surface profilometric measurement of a 3-D object," *Proc. Of the Eighth Asia-Pacific Conf. on Non-destructive Testing*, Taipei, Taiwan, 1995, pp. 966-974.
5. J.A. Jalkio, R.C. Kim and S.K. Case, "Three dimensional inspection using multistripe structured light," *Optical Engineering*, Vol. 24, No. 6, 1985, pp. 966-974.
6. M. Magee and R. Weniger, "Isolation of three dimensional features of known height using a light-stripped stereoscopic system," *SPIE*, Vol. 1822, 1992, pp. 175-187.
7. Y.F. Wang, A. Mitichi and J.K. Aggarwal, "Computation of Surface Orientation and Structure of Objects Using Grid Coding," *IEEE*, 1987

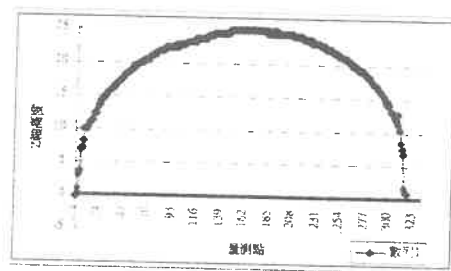
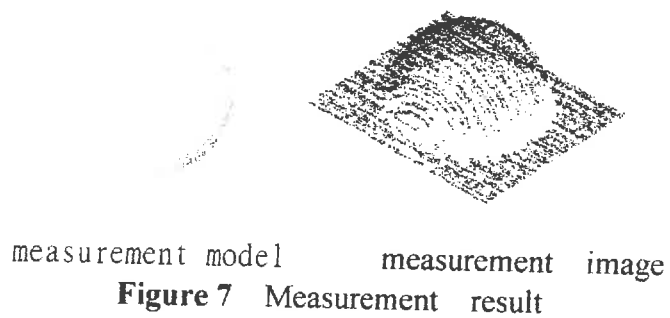


Table 1 Calibration result of the semi-spherical ball

translation (mm)	actual (mm)	error (mm)
90.3	88.8	1.5
68.7	68.9	-0.1
52.8	53.8	-0.9
47.7	48.8	-1.0
34.0	33.8	0.2
24.4	23.9	0.5
84.5	83.9	0.6
73.6	73.9	-0.2
57.9	58.8	-0.8
43.3	43.9	-0.5
33.8	33.9	0.0
24.8	23.8	1.0
84.8	83.9	0.9
73.2	73.8	-0.5
62.9	63.8	-0.8
47.8	48.9	-1.0
38.4	38.8	-0.3
24.8	23.9	0.9
84.9	83.9	1.0
74.0	73.9	0.1
62.8	63.8	-0.9
52.9	53.8	-0.8
38.5	38.9	-0.3
25.3	23.9	1.4

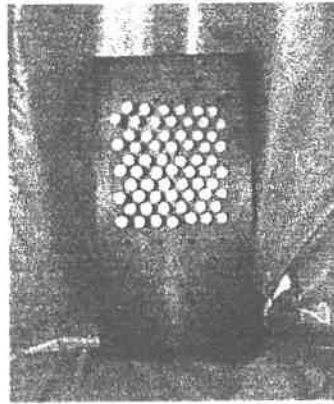


Figure 4 CCD calibration plate

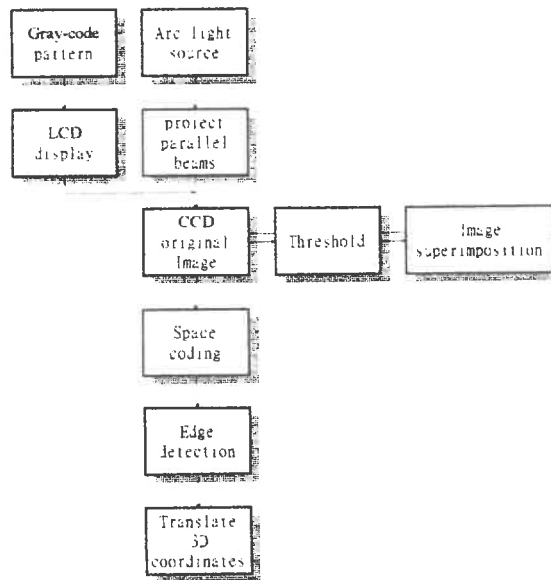


Figure 5 Block diagram for 3D measurement

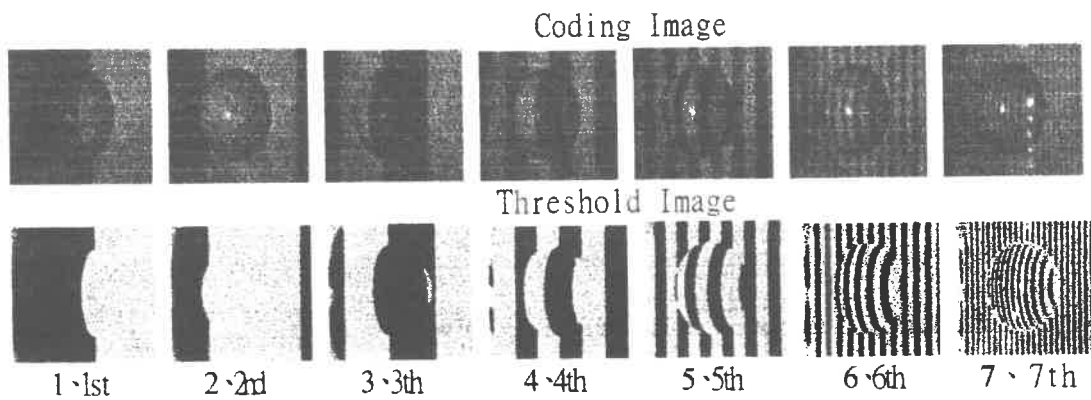


Figure 6 Coding and threshold images

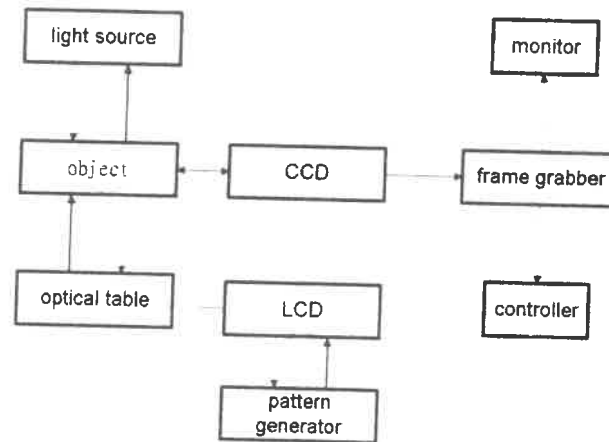


Figure 1 System block diagram

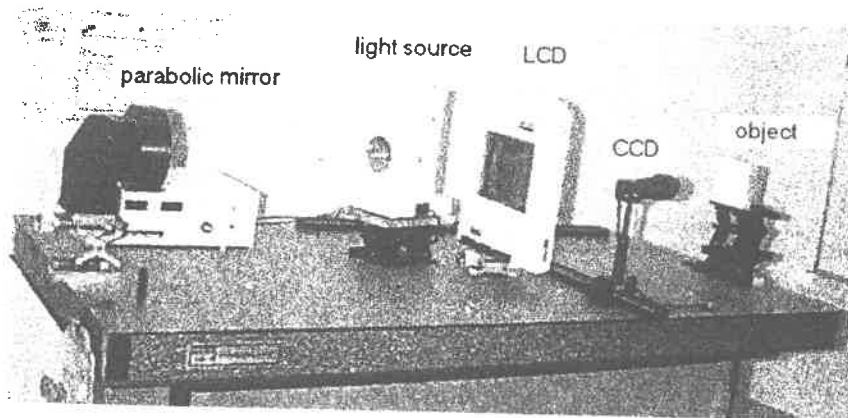


Figure 2 A view of the measurement system

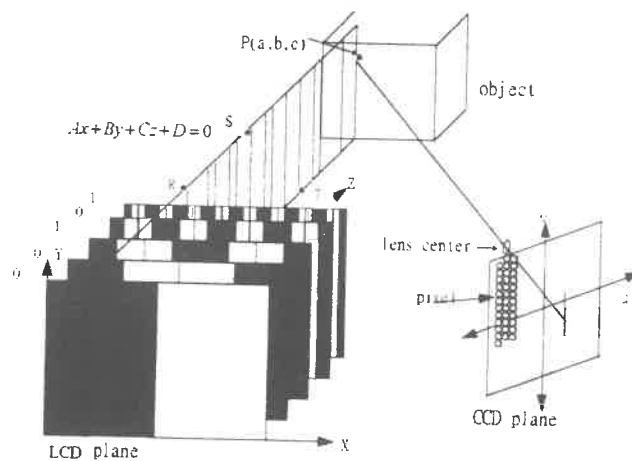


Figure 3 Geometry used for derivation of relation between space code and CCD focus plane

A METHOD TO ANALYZE ZONE-PLATE INTERFEROGRAMS BY USING A NEMATIC LIQUID-CRYSTAL

Takashi Nomura¹, Kazuhide Kamiya¹, Hiroshi Miyashiro¹,
Kazuo Yoshikawa², Hatsuzo Tashiro², Hiroyoshi Onnagawa², and Hiroyuki Okada²

1. Toyama Prefectural University, Kosugi-machi, Toyama, 939-03, Japan

2. Toyama University, Gofuku, Toyama, 930, Japan

Interferometers are used to measure errors in the shape of mirror surfaces. They are classified into two groups, common-path interferometers and non-common-path interferometers. In common-path interferometers, the measurement and reference paths pass through essentially the same space and therefore the interference fringes are scarcely affected by air turbulence in the optical paths. The zone-plate interferometers proposed by Smartt¹ belong to the group of common-path interferometers. The zone plate used in the interferometer is manufactured to generate a diffracted wavefront which corresponds to the design shape of the mirror surface. When a mirror surface under test has a slight error in its shape with respect to the design surface, interference fringes are observed on the image plane of the interferometer. Taking advantage of this feature, we developed a small type of zone-plate interferometer that could be mounted on an ultra-precision lathe². An error in the shape of the mirror surfaces manufactured by the lathe is measured by analyzing the interference fringes. A phase-shifting interferometry (PSI)³ measurement, in which a reference mirror attached to a piezoelectric transducer is used as a phase shifter, cannot be used to analyze the fringes, because the reference point is located on the mirror surface under test. Therefore we used the Fast Fourier Transform method in which the fringes with a carrier component were analyzed.⁴ However this method of an analysis needs much time. To finish the analysis of the zone-plate interferograms in a short time, a PSI device for the zone-plate interferometer is needed. Nematic liquid crystal devices have been developed for displays such as those for liquid crystal television sets. The nematic liquid crystal devices have the property of modulating the phase as well as the intensity of light. In the zone-plate interferometer measurement the liquid crystal cell is used as a phase modulator. Our proposal is to use liquid crystal as a zone-plate PSI device.

The zone plate is a computer generated hologram and generates a wavefront whose shape is identical to the design shape of a concave mirror surface. The design mirror surface is symmetrical when rotated about its optical axis. The zone plate is illuminated by a laser source converging at the vertex O of the mirror surface as shown in Fig. 1. The zone plate is made in such a way that one of the first orders of diffracted light illuminates the entire surface of the mirror. We denote the first order of the diverging diffracted light as +1. The diffracted light +1 generates a wavefront which impinges on the mirror in a direction perpendicular to the design surface. The light is reflected by

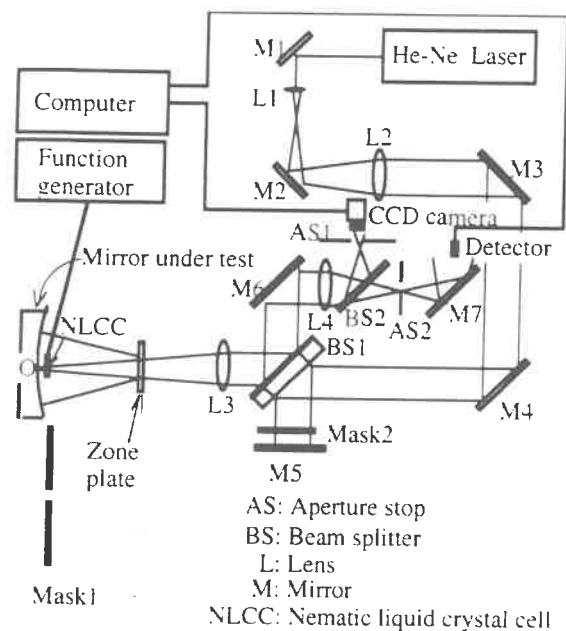


Fig. 1 Configuration of a zone-plate interferometer

the mirror surface under test and returns to the zone plate, where it is again diffracted and then recollimated by lens L3. The wavefront of the diffracted light is affected by any error in the shape of the mirror under test and, therefore, the diffracted light of order $(+1,+1)$ is used as the test light. The diffracted light of order $(0,0)$, actually non diffracted light, generates a spherical wavefront coming from the vertex of the mirror surface under test. The wavefront is not affected by any error in the shape, and is used as the reference light. The light beams of order $(+1,+1)$ and $(0,0)$ go through lens L3. Because the wavefront passing from lens L3 to lens L4 is plane, it is focused into small spot at aperture stops AS1 and AS2. Extraneous

wavefronts are blocked out almost completely by the aperture stops. Clear interference fringes are obtained on the image planes of the charge coupled device (CCD) camera and detector.

The nematic liquid crystal cell is located in front of the vertex of the mirror under test in this method. An A.C. voltage is applied to the liquid crystal cell. Only the phase of the light passing through the nematic liquid crystal cell is varied by the change in voltage. The intensities of the zone-plate interference fringes vary according to the phase difference between the test light and the reference light. The phase difference is, therefore, obtained by measuring the intensities of some of the fringes. The phase difference is determined by errors in the shape of the mirror. The four-step algorithm⁵ is used to analyze the phase difference. The intensities of the four interferograms are measured by the CCD camera and stored in the computer memory. The phase difference is obtained from the four intensity patterns and can be easily related to the error in the shape of the mirror under test.

Zone-plate gratings with binary phase structures were formed on one surface of the glass plates by a thin SiO_2 film. The other surface was coated by SiO_2 to prevent reflection. The intensity of the test light should be equal to that of the reference light to obtain interference fringes with a high visibility. The zone plates were, therefore, made in such a way that the first order diffraction efficiency of the zone plates was equal to the zero order diffraction efficiency. The following homogeneous type of nematic liquid crystal cell was made: liquid crystal material with a diameter of 3.0mm and a thickness of $27\text{ }\mu\text{m}$ was inserted between two plates of glass with dimensions of $10\text{mm} \times 10\text{mm} \times 1\text{mm}$. Because the time response of the phase modulation is slow and the

measurement of the fringe intensities requires a long time when the nematic crystal cell is used, we used the following technique: The function generator generates a square-wave signal with a fixed carrier frequency and a saw-tooth amplitude. The generator signal goes to the liquid crystal cell. The carrier frequency was 1 kHz, the frequency of the saw-tooth signal was 0.15 Hz, and the amplitude of the saw-tooth signal varied from 1.5 V to 2.5 V. The phase of the reference light changes according to the amplitude. The reference light and the test light go to beam splitter BS2 and are divided into two directions. The light reflected by the beam splitter passes through aperture stop AS1, and received by the CCD camera, where interference fringes are produced in which the intensity of the interferograms changes continuously. The rays of light transmitted through the beam splitter pass through aperture stop AS2, and are received by the detector, where similar interference fringes are produced. The detector continuously measures the intensity at one point of the interference fringes. Because the amplitude of the signal is varied periodically, the intensity of the interference fringes also varies periodically. The intensity is measured by the computer which determines. The maximum, the average, the minimum and the next average value of the intensity

signal from the detector. When the detector signal is equal to one of these values, the computer gives an order to the control units and the interferogram intensities of the entire mirror surface produced by the CCD camera are stored in a frame of the memory of the computer. The phase difference is analyzed from the intensities.

A spherical concave mirror with a radius of curvature of 500 mm and a diameter of 100 mm was measured by this method. The photograph of the experimental setup is shown in Fig. 2. Figure 3 shows a zone-plate interferogram, taken when the liquid crystal cell was not placed in front of the mirror under test. Figures 4 (a), 4 (b), 4 (c), and 4 (d) show the zone-plate interferograms obtained by this method, in which the phase angle was shifted in steps of $\pi/2$ each interferograms. The phase distribution analyzed by the PSI is shown in Fig. 5.

The tests showed that in zone-plate interferometry a nematic liquid crystal cell can be used as a phase modulator. The phase distribution of the error in the shape of a concave mirror could

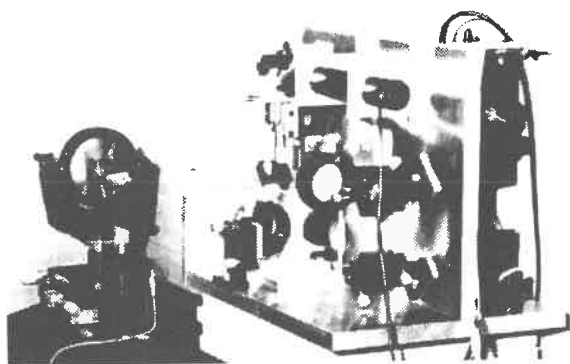


Fig. 2 Photograph of experimental setup

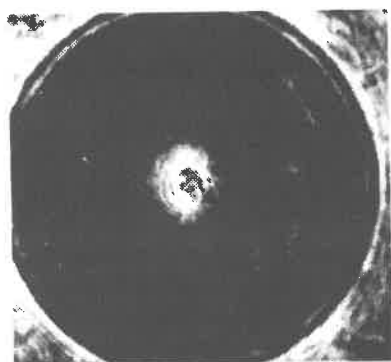


Fig. 3 Interferogram obtained by the zone-plate interferometer. The liquid crystal cell was not placed in front of the mirror under test.

be analyzed by phase-shifting interferometry using a liquid crystal cell.

Acknowledgments

We would like to thank Chisso Petrochemical Corp. for supplying the liquid crystal material. We also wish to thank Mr. F. Kobayashi of Fuji Photo Optical Co., Ltd. and Mr. T. Shimizu of Nachi-Fujikoshi Co., Ltd. for their support.

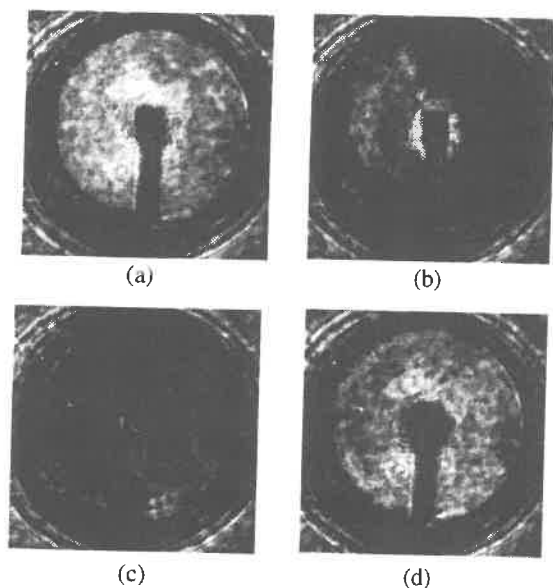


Fig. 4 Interferograms obtained by the zone-plate interferometer. The liquid crystal cell was placed in front of the mirror under test. (a) Phase modulation of 0, (b) Phase modulation of $\pi/2$, (c) Phase modulation of π , and (d) Phase modulation of $3\pi/2$.

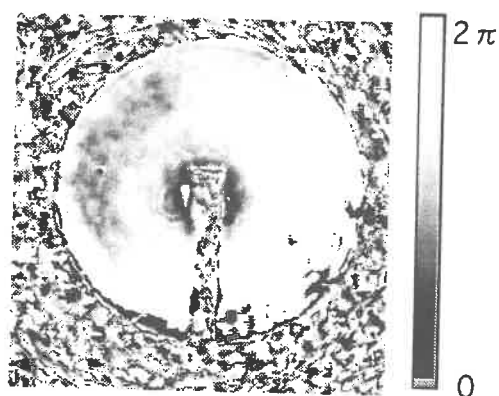


Fig. 5 Phase distribution analyzed by the phase shift interferometry

References

1. Smartt, R.N. "Zone plate interferometer," *Appl Opt*, 1974, 13, 1093-1099
2. Nomura, T., Yoshikawa, K., Tashiro, H., Suzuki, M., Kobayashi, F., and Usuki, M. "Shape measurement of workpiece surface with zone-plate interferometer during machine running," *Prec Eng*, 1993, 15, No.2, 86-92
3. Burning, J.H., Heriot, D.H., Gallagher, J.E., Posenfield, D.P., White, A.D., and Brangaccio, D.J. "Digital wavefront measuring interferometer for testing optical surface and lens," *Appl Opt*, 1974, 13, 2693-2703
4. Kamiya, K., Nomura, T., Miyashiro, H., Yoshikawa, K., Tashiro, H., Ozono, S., and Suzuki, M. "An analysis of zone-plate interference fringe pattern," *Proceedings of international symposium on optical fabrication, testing, and surface evaluation*. Tokyo, 1992, 1720, 311-316
5. Malacara, D. *Optical shop testing*, New York, Wiley, 1992, 501-513

Key Words: zone plate interferometer, nematic liquid crystal, phase shifting interferometry

Straightedge Metrology in a Wet Environment

Scott B. Peirce, Daniel B. DeBra
Department of Aeronautics and Astronautics
Stanford University, Stanford, CA

Abstract

Thermal control and uniformity of temperature in machine tools and metrology instruments is a fundamental limitation to precision engineering. The problem scales unfavorably with larger devices. It is well established that oil showering provides the best means of thermal control by at least an order of magnitude. In the Quiet Hydraulics lab, we have developed a family of laminar flow based hydraulic actuators for ultraprecision applications: a laminar flow spindle motor, a fast, short-stroke linear actuator, and hydrostatic carriage ways with built-in provision to correct for straightness and yaw error motions. These are compatible with an oil shower and offer improved thermal performance since the working fluid carries away the heat of inefficiency. Hydraulic sensing and actuation offer a low cost, high performance precision alternative to optical and electrical sensing and actuation techniques.

In order to demonstrate clearly the benefits of this technology and enhance its transfer, we are completing the integration of these individual components into an operational precision machine tool. The remaining tasks to achieve this goal are: the development of the metrology and compensation for non-straightness and yaw motions of the carriage ways, motion control of the carriage ways, and the integration of these technologies into a fully functional machine.

This paper describes a new method of non-contacting straightedge detection, called a "differential oil gauge." We are integrating the technique into our Quiet Hydraulics testbed to correct for motion of the carriages. The actuation is provided by slideway bearings which modulate their gap as the means of non-straightness and yaw correction. We will provide a description of the current hardware, the theoretical basis for its design and the status on its implementation in our testbed.

Differential Oil Gauge Design

Since no slideway is perfectly straight, there must be provision for measuring and correcting error motion if additional accuracy is needed. Typical straightedge measurements require resolutions on the order of 10 nm over a span of 25 μm . LVDT's and interferometers are frequently used. In the implementation of an LVDT based straightedge sensor, one must ensure that the probe will not lift off the straightedge surface. Contamination between probe and surface can give spurious readings. Care must be taken to ensure that the probe always traces the same line.

	Hydraulic	Capacitive	LVDT	Optical	Eddy Current
Meets current accuracy requirement of 10nm?	Yes	Yes	?	Yes	?
Contamination sensitivity	Low	Medium	High	High	?
Repair Technology Needed by Machinist	Mechanical OK	Electrical ?	Electrical ?	Optical ?	Electrical ?
Approx. Cost of readout	\$50	\$5,000	\$1,500	\$6,000	\$5,000
Thermal environment	Best	Best	Good	(isolated)	Poor ?
Material Requirements and Concerns	No special requirements Lap is OK, no polishing required	Conductive	Hardness	Coating or polish required	Conductor, no electrical runout
Contacting	No	No	Yes	No	No

Table 1: Comparison of Candidate Straightness and Yaw Sensors

We are developing and have built a novel hydraulic sensor for measuring non-straightness and yaw called a "differential oil gauge". The sensor can be thought of as a pair of opposed externally pressurized bearings. There is an external resistance between the supply and the pocket and a variable resistance to flow between the lands and the surface of the straightedge. The resistance is a function of the gap, h , between the lands and the surface of the straightedge. By measuring the differential pressure between the two pockets over a small range of motion, one can arrive at a linearized relationship with the deviation of the head from the nominal centered position (See Figure 1). A differential oil gauge measurement has no probe contact: There is neither lift off nor surface contamination. Furthermore, since the oil gauge measurement averages over the straightedge surface, sensitivity to errors with short correlation distance is reduced.

Initial evaluations of a prototype oil gauge sensor are very promising. This technology has the potential of being both more robust and less expensive than the use of LVDT's or optical devices as straightedge sensors in a oil shower environment. The performance is consistent with the theoretical model on which the design is based. It has demonstrated

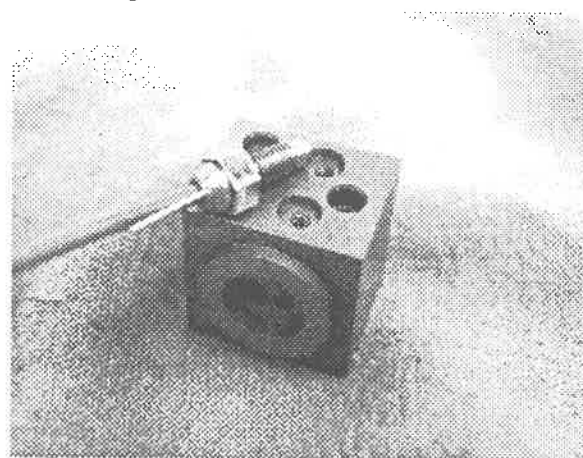
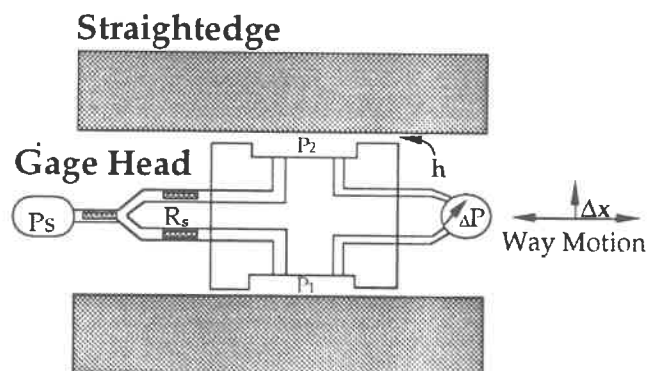


Figure 1: Oil Gauge Design

Method	Advantages	Disadvantages
Area Reduction	Simple solution	Slight reduction in stiffness
Force Compensation	Dramatically reduce stiffness	Requires precision hydraulic resistance networks
Force Calibration	Account for effect of forces	Complicated external calibration
Active Nulling w/Valve	Zero effective stiffness	Requires costly flapper valve
Active Nulling w/Servo	Zero effective stiffness	Requires external servo system

Table 2: Relative Merits of Proposed Methods for Stiffness Reduction

a good signal to noise ratio and adequate linearity, but exhibits an undesirable stiffness. Although it is possible with brute force designs to make the machine tool extremely stiff, this is not an effective solution. We have five approaches to reduce or eliminate the stiffness which we are continuing to evaluate (Table 2). The first is simply to reduce the area of the pockets. The second is to include a model of the net disturbance force in the straightedge reversal process. The third approach is a redesign of the oil gage sensor to minimize the net disturbance force by crossfeed to a second set of pockets (force compensation). The fourth alternative is to implement a feedback control system to actively null the pressure differential by actuation of a flapper valve, which regulates the two flow resistances comprising the outer arm of the hydraulic bridge. The fifth will be to servo the position of the sensor head so it always stays in a position of zero pressure differential. The last two methods are more complicated but have the added advantage of being independent of the variation in supply pressure.

Straightness and Yaw Feedback

The straightness measurement system consists of a pair of differential oil gauges and a reference straightedge (Figure 2). The common-mode signal from the oil gauges provides information about the non-straightness and the differential signal provides information about the yaw in the motion of the carriage. This information will be used in conjunction with data from optical scales which will measure the long-stroke motion of the carriage. The reference straightedge is imperfect and must be mapped out using reversal. These data are then used in the form of a look-up table to correct the measurements as a function of the carriage position along the ways.

The feedback control of straightness and yaw is the next concern to be handled. A high bandwidth dedicated analog controller will command the actuators. The straightedge correction command input to this regulator will come from a digital controller. The controller uses the carriage position along the ways as entry to the look-up tables. The look-up tables contain a map of the irregularities of the straightedge. External disturbances will be compensated for by the analog controller. The look-up table corrections will be

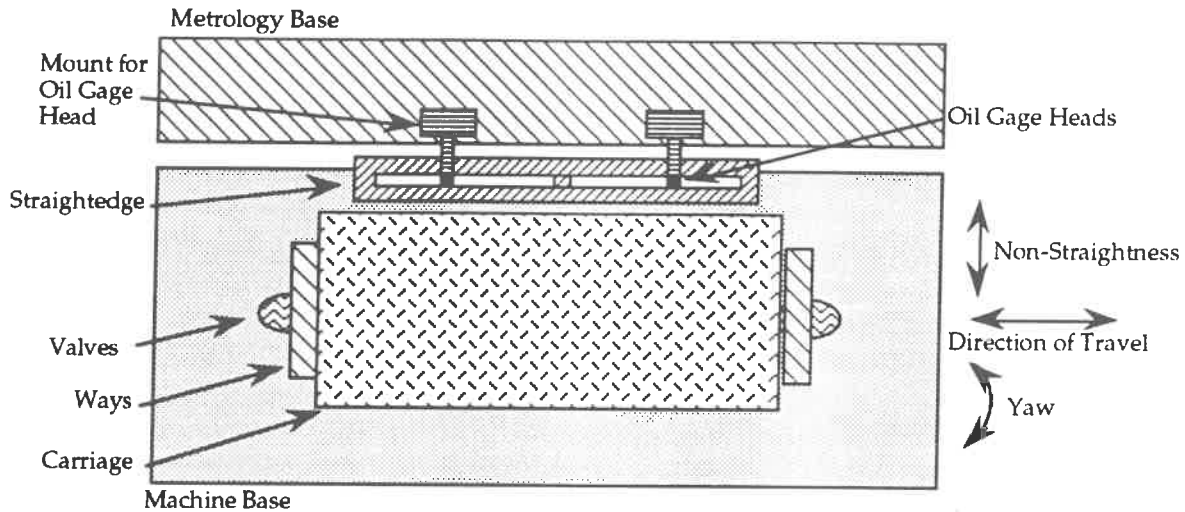


Figure 2: Straightness and Yaw System Schematic

handled by the lower bandwidth digital controller, which is adequate for the relatively slow feed rates required in precision turning or grinding.

This design is consistent with a fail-operate/fail-safe philosophy. Any failure would still permit the continued operation of the carriage, although in a degraded state. Without the look-up table corrections, the system performance degrades to the errors inherent in the straightedge, and if the servos fail, the errors then become the errors in the carriage ways themselves. There are advantages in correcting the error in the ways: when fully operational, it presents a "perfect" machine tool to the machine tool controller. The controller is simpler and one can then utilize a commercial unit without modification.

Acknowledgements

We would like to acknowledge the support of the National Science Foundation (both prior (NSF-DDM-8914232) and current support).

References

- Bryan, J. B. . International Status of Thermal Error Research. *Keynote Paper, Annals of the CIRP Vol. 39/2*. 1990.
- DeBra, D. et al., Ultra Precision Machining: Quiet Hydraulics and Computer Support Research. Final project report NSF-DDM-8914232, "Strategic Manufacturing Initiative: Ultra Precision Machining." 1994.
- Donaldson, R., A Simple Method for Separating Spindle Error from Test Ball Roundness Error. In *CIRP Annals Vol. 21*, 1972.
- Peirce, S., Kahn, R, and Rischer D., *Design of an Oil Gage*. ME119 Class Project, Stanford University, 1992.
- Peirce, S., Tran, H., Wiedemann, M., and DeBra, D., "Quiet Hydraulics for Ultraprecision Actuation," ASPE 1994 Spring Topical Meeting, Tucson Az.

MACHINING VIRTUAL TEST PARTS

by

Tony Schmitz, John Ziegert
University of Florida
Gainesville, FL

Introduction

The current procedure followed to manufacture a new part by CNC machining is to write the part program, machine a test part and measure the test part for conformance to the required dimensions and tolerances. If the test part dimensions are not correct, the part program is modified and the process repeated until a successful part is machined. In many applications, such as the aerospace industry, where material cost and machining time are high, this iterative process becomes economically unacceptable.

Research has been conducted to test the feasibility of using the Laser Ball Bar (LBB), a spatial coordinate measuring device, to measure dynamic continuous-path contours to micrometer accuracy prior to machining. In this way, a virtual test part can be measured and compared to the design drawings to validate the CNC part program. This reduces or eliminates the costly and time-consuming steps involved in the machining of physical test parts.

Laser Ball Bar

The LBB is a precision linear displacement measuring device developed by Ziegert and Mize at the University of Florida [1]. It consists of a two-stage telescoping tube with a precision sphere mounted at each end. A heterodyne displacement measuring interferometer is aligned inside the telescoping tube and measures the relative displacement between the two spheres. See Figure 1. The LBB has been shown to be accurate to sub-micrometer levels during static measurements of spatial coordinates [2].

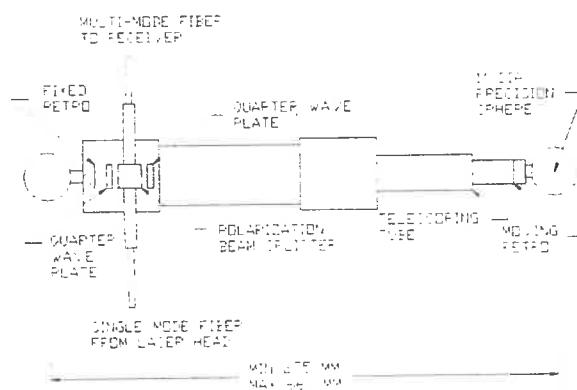


Figure 1: Laser Ball Bar

Once initialized, the LBB uses trilateration to measure the spatial coordinates of points along a CNC part path. The six sides of the tetrahedron formed by three base sockets (attached to the machine table) and a tool socket (mounted in the spindle) are measured and, by geometry, the coordinates of the tool position can be calculated. In sequential trilateration, the same part

path is traversed three times, measuring the lengths of one of the base-to-tool socket legs at a finite number of points during each repetition. See Figure 2. In simultaneous trilateration, three LBBs ride on a single sphere at the tool point to completely define the tetrahedron with one execution of the CNC program. See Figure 3.

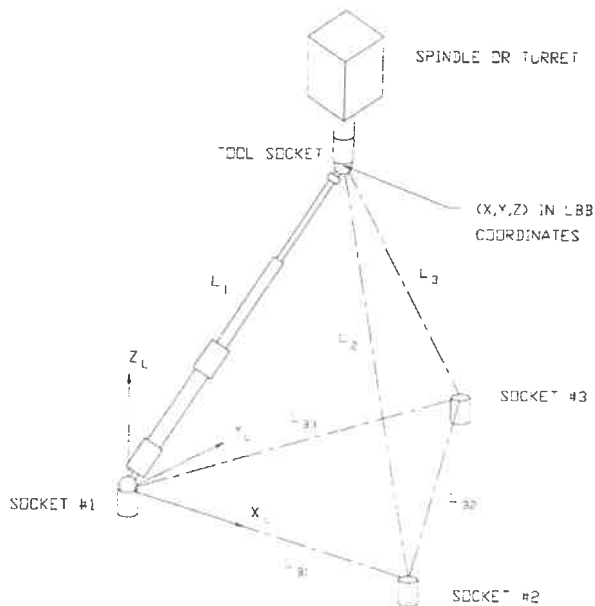


Figure 2: Sequential Trilateration

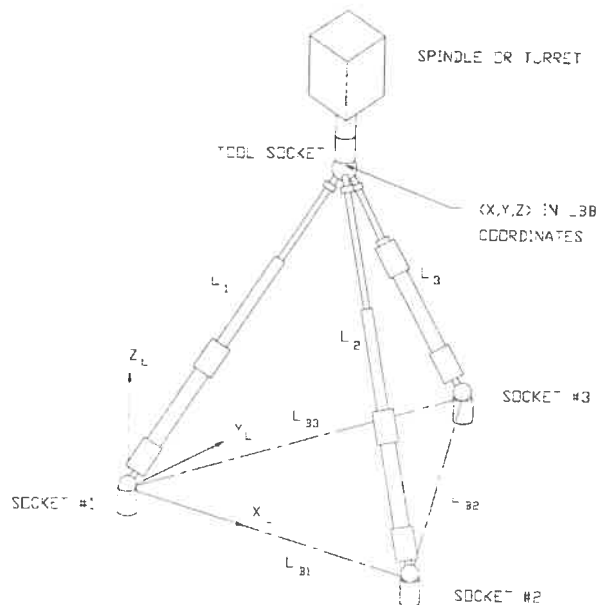


Figure 3: Simultaneous Trilateration

Dynamic Path Measurements

The purpose of this research was to evaluate the dynamic measurement performance of the LBB and the feasibility of using sequential trilateration to perform these measurements. In other words, it was desired to measure the spatial coordinates of the tool point at predefined intervals along the contour during the execution of the CNC part program while the machine was in motion using one LBB. These dynamic path measurements using a single LBB require that the three leg lengths from the base sockets to the tool socket are acquired as the tool passes through the same points along its path on three consecutive runs. This necessitates a repeatable triggering method to initiate data capture.

Two separate techniques were implemented and tested to provide the repeatable measurement trigger. The first sampling procedure used a time-based scheme. In this method, a sample was taken at the start of the motion and then consecutive samples were taken at well-defined time intervals (every 5 ms) throughout the part path. This method was shown to be unacceptable, since the acceleration profile of the controller varied slightly from run to run on the two-axis turning center used in this study. This caused the tool to be in a slightly different position along the path during each leg measurement.

The second triggering scheme involved the use of the feedback signals from the machine axis encoders. In this method, the axis encoders were continuously sampled and the length of the LBB recorded at discrete distance intervals along the part path. This sampling method was chosen because the positioning repeatability (and therefore, the encoder output) is the smallest source of error in the contouring performance of a machine tool.

Experimental Results

Initial dynamic path measurements using the X and Z-axis encoder sampling algorithm were completed for both right-angle and semi-circular contours. Experimental results showed general path degradation with higher velocities. This path degradation was especially evident for the right-angle contour. At velocities of 1 m/min and higher, the corner was increasingly rounded due to a feature in the controller which rounds corners to avoid overshoot.

Since the actual part path changed as the feed rate was increased, it was impossible to distinguish between path errors caused by the controller and measurement errors introduced by the LBB. Therefore, in order to provide a contour which would be independent of velocity, the spatial coordinates of a circular contour provided by the spindle rotation and a socket mounted off-center were measured by sequential trilateration. The circular path traced by the socket as the spindle rotated was divided into 256 parts. An LBB data point was taken when the angular position of the spindle (monitored by reading the spindle encoder) was an integer multiple of the sample resolution. The path velocity (spindle speed) was then varied to investigate the dynamic measurement performance of the LBB. The radial error motions for a 60 rpm test are shown in Figure 4.

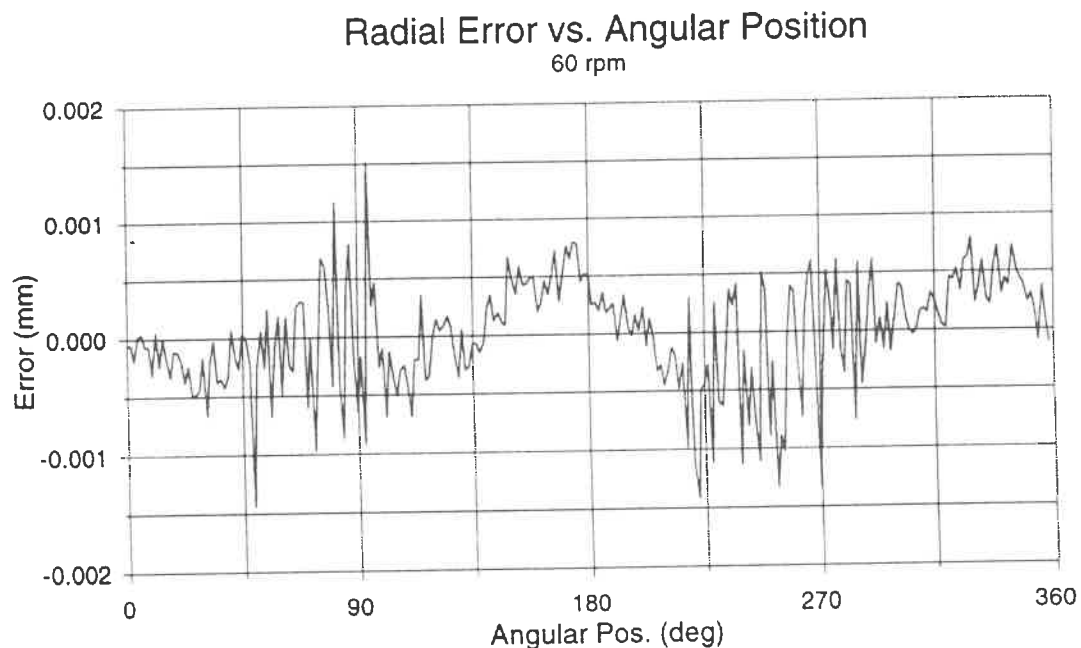


Figure 4: Radial Error Motion Using LBB

In order to validate the LBB results, the spindle error motions were measured at the same feed rates using two mutually-perpendicular capacitance probes and a precision artifact [3] and compared to the radial error motions attained by the LBB tests. See Figure 5. It was then concluded that the LBB is suitable for dynamic contour measurements since the magnitude of the radial error motions of the circular path as measured by the LBB corresponded to the spindle error motions measured with the capacitance probes.

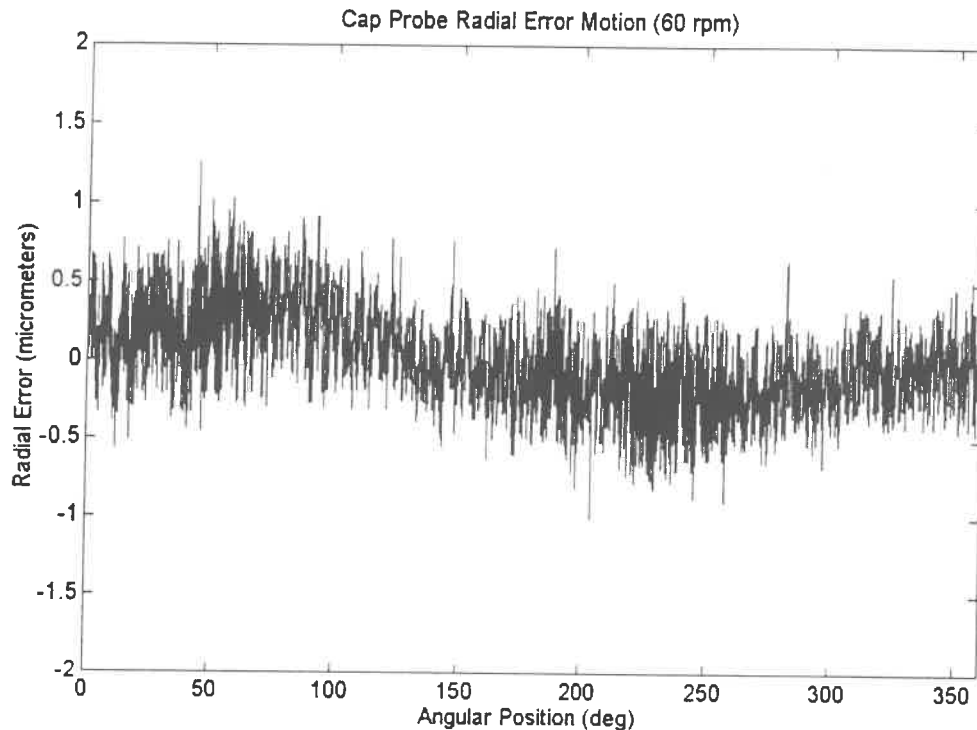


Figure 5: Spindle Error Motions Using Capacitance Probes

Conclusions

Experimental results show that the LBB is a suitable device for dynamic path measurements of three-dimensional CNC tool paths, or virtual test parts. These virtual test parts can then be used to validate the conformance of the CNC part program to the required engineering specifications. Future work includes modifying the LBB design to allow simultaneous trilateration using three LBBs and adding a cutting force model to the measurement algorithm to account for deflections during machining.

1. Ziegert, J.C. and Mize, C.D. "Laser Ball Bar: A Precision Instrument for Machine Tool Metrology," *Precision Engineering*, Vol. 16/4, October 1994, pp. 259-267.
2. Mize, C.D., Ziegert, J.C., Pardue, R. and Zucker, N. "Spatial Measurement Accuracy Tests of the Laser Ball Bar," Final Report for CRADA No. Y-1293-02244 between Martin Marietta Energy Systems and Tetra Precision, Inc., August 1994.
3. Tlustý, J. "System and Methods of Testing Machine Tools," *Microtecnic*, Vol. XIII, 1959, pp. 166-167.

A Microfabricated Magnetic Field Sensor Based on Electrodeposited Giant Magnetoresistant Material

W. Wang, A. Natarajan, E. Ma and C. Khan-Malek

**Microsystems Engineering Team, Department of Mechanical Engineering,
Louisiana State University, Baton Rouge, LA 70803**

I. Introduction

Magnetic field sensors are used in a wide range of measurement situations, including those encountered in aerospace endeavors. The large variety of magnetic sensors currently available are mainly based on two mechanisms: the Hall effect, and the semiconductor magnetoresistivity [1]. Most of micro magnetic field sensors are fabricated with the silicon technology. These sensors offer some obvious advantages: microsize, low cost resulting from batch production technology, as well as the convenience for integration with signal processing circuit. Efforts have been ongoing to achieve increasingly higher resolution. For example, Kadar et al.[2] proposed a magnetic-field sensor based on a resonating single-crystal silicon structure. The excitation of the resonator is achieved by the Lorentz force generated by a sinusoidal current flowing through a rectangular coil deposited on the structure. The amplitude of the vibration, which is proportional to magnetic field, is measured by sensing capacitors. The high quality factor of the resonator makes it possible to achieve high resolution. Kawahito et al. [3] proposed a combination of ferromagnetic (soft magnetic) materials and silicon technology to take the advantages of both materials. They reported a fluxgate magnetic sensor composed of micro-structured solinoids for the excitation and pick-up coils and permalloy cores formed by electroplating. It is claimed that relatively high sensitivity and low-offset characteristics have been achieved compared with those of the conventional silicon magnetic sensors and its resolution and sensitivity make it suitable for application to magnetic compass.

Compared with the Hall effect magnetic sensors, the magnetic sensors based on magnetoresistance effect have some advantages. The classical technique of Wheatstone bridge technique as well as modulation technique can be used to process the otherwise weak signal, resulting a much higher measurement accuracy. Since the late 80s, Giant Magnetoresistance (GMR) effect in magnetic multilayer thin films has attracted considerable attention [4, 5, 6, 7]. $\Delta R/R$ values in the range of 5-150% (depending on material and temperature) have been reported for a number of multilayers composed of ferromagnetic - antiferromagnetic couples. These multilayers were initially prepared using physical vapor deposition techniques, such as sputtering and molecular beam epitaxy. A little over two years ago, it was demonstrated that GMR multilayers can also be prepared using electrodeposition in a single plating bath by a careful control of bath compositions and cyclic variation of plating potentials. This technique offers an inexpensive route for the preparation of "thick" (on the order of many microns) GMR multilayers. In addition, it allows us to incorporate this group of novel materials into microfabrication for the first time with the fast developing Microelectromechanical Systems (MEMS) technology.

To date, little effort has been spent to exploit this new group of GMR materials in applications. The magnetic field dependence of GMR before saturation (which typically occurs in the range of 0.05 to 5 Tesla depending on the material system) makes it possible to pursue their application as field sensors. It has been observed that this dependence is very strong, monotonic, and almost linear in a wide range of magnetic field. When measuring relatively low fields or small field changes (small $\Delta R/R$), it is useful to increase R values for improved sensitivity and accuracy. As will be demonstrated in this work, this can be achieved by lithographically patterning the blanket multilayer film without increasing the size of the sensor. The design, analysis, and some

preliminary experimental results of a GMR wire microsensor is presented in this paper. Our work demonstrates that magnetic field sensors based on microfabrication and GMR material can be fabricated which: (1) are micro-sized and batch fabricated with MEMS technologies, (2) have the desired property of giant magnetoresistance, (3) have relatively high sensitivity in a range of magnetic fields, and (4) have stable and predictable measurement accuracy and resolution. The device can also be used to study the property of the GMR materials which are still not fully understood when the magnetic field is weak.

II. Implementation Strategy/Fabrication Method

Design and Analysis

The design of the sensor is illustrated schematically in Figure 1. Generally speaking, to have a better sensitivity, a larger magnetoresistance is always desired. The total length, the width, and the thickness of the GMR wire required are determined by the desired total magnetoresistance. The ohmic resistance of a resistive wire is a function of resistivity ρ , the total length L , and the cross-sectional area A , in the form:

$$R = \rho L / A.$$

Assuming a constant resistivity ρ , we either need to increase L or decrease A to have the largest possible value of R . In our design, the microfabrication technique provides the geometric freedom (e.g., a zigzag pattern) to reach the desired large value of magnetoresistance (ΔR) in a micro-sized device.

The magnetic field sensor is designed to occupy a planar area of 1.0 millimeter by 1.0 millimeter. The GMR wire has a cross-sectional area of 3 micrometers by 3 micrometers. The gap between strips is designed to be 3 μm . The total length of the wire is tentatively designed to be roughly equal to 150 mm. Therefore if the magnetoresistivity of the GMR material is about $1 \times 10^{-6} \Omega \mu\text{m}/\text{Gauss}$ at room temperature [5], the total magnetoresistance of the proposed magnetic field sensor would be $2 \times 10^{-2} \Omega/\text{G}$. This resistance change is readily measurable, which makes the sensor usable even for small changes in magnetic field. An optical mask for lithographic fabrication has been designed and made based on this design.

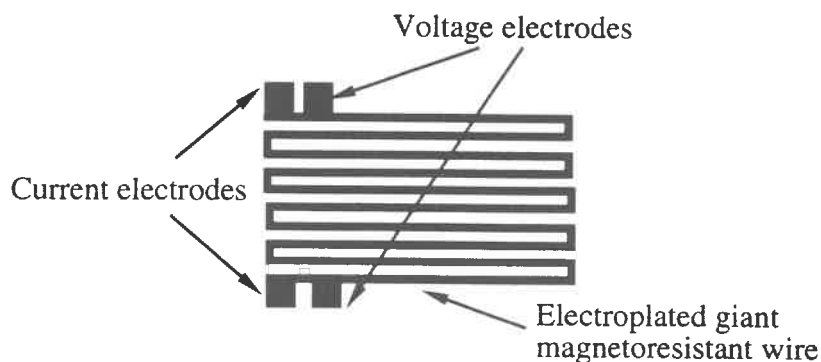


Figure 1. Schematic design of meandered shaped magnetic field sensor based on giant magnetoresistive effect

Fabrication

The fabrication of this sensor is being carried out in collaboration with the LSU Center for Advanced Microstructures and Devices (CAMD) which houses a synchrotron light source

dedicated primarily to lithography, and a Class 100 cleanroom with all necessary equipment for micromachining activities. The electrodeposition processing is being conducted with computer-controlled electrodeposition systems in the Microsystems Engineering Team (μ SET) Laboratory in the Mechanical Engineering Department at LSU.

The fabrication steps flow as follows. An oxidized Silicon wafer is used as the substrate. The substrate is coated with Copper (1nm) as the plating base. After deposition, the blanket multilayer film is patterned using standard lithography techniques. An optical resist of thickness, equivalent to the thickness of the magnetoresistive wire is applied over this copper base. It is then exposed and developed to generate the desired magnetoresistive wire pattern. The giant magnetoresistive material is then electrodeposited into the micro-grooves using pulsed potential deposition (see below). After deposition, the remaining resist is stripped followed by stripping of the copper underneath. Finally, contact pads are generated by electrodeposition of copper at the two ends of the wire. A mask with the desired feature has been manufactured as shown schematically in Figure 1. The pattern generated as such consists of "long wires" to increase the total resistance. The GMR in micro features will be calibrated using standard four electrodes techniques. Magnetic field will be generated using precisely fabricated coil and accurate current control.

Specifically, we started this project with the Co-Cu multilayers. The electrodeposition of multilayers followed the procedures reported before by other groups. An oxidized silicon wafer was used as substrate. A very thin Cu layer, a couple of hundred angstrom in thickness, was sputter-deposited (or evaporated) onto the wafer as plating base. The use of such thin plating base is to ensure that the resistance of this layer is very large so that it will not interfere with the resistance measurement of the thick (~ 3 -5 microns) GMR multilayer.

The Cu/Co multilayer was deposited into the microfeatures shown in Fig. 1, after pattern generation (Fig. 1). The electrochemical deposition was made from a single sulfate bath containing both the Cu and Co ions, with a concentration ratio $<1:10$, and boric acid. Cu and Co were deposited at -0.2 and -1.0 V, respectively. At -0.2 V, only Cu is deposited, whereas at -1.0 V, both metals are co-deposited. Because the Cu concentration in the electrolyte is very low, the deposition of Cu is diffusion limited. Therefore, with a fast Co deposition, only a small amount of Cu ($<10\%$ as determined by energy dispersive X-ray (EDX) analysis) is incorporated in the Co layer. During deposition, the applied potential was controlled by a potentiostat which alternates the deposition potentials. The multilayer deposited as such consists of a large number of alternating nanolayers (5-10 nm) of Cu and Co (e.g., [Cu 2nm/Co 5 nm]₄₀₀) and the total thickness is about 3 microns. Typically, the entire deposition process takes about 9 hrs.

The deposits grow as such from a polycrystalline plating base. This is different from the majority of previous work, where epitaxial growth from a single-crystal substrate is often used. However, as already demonstrated by several groups, multilayers electrodeposited in micro-sized features from polycrystalline plating base still exhibit decent GMR behavior. We therefore anticipate that GMR exhibited by these multilayers will suffice for our applications. The GMR effect of multilayers deposited this way is being compared with literature data for epitaxial multilayers. It should also be noted that our systematic tests later in the project will also reveal the GMR behavior at very low magnetic fields, which is difficult to measure in macro-scaled films.

Calibration of the Proposed Microsensor

The LIGA-fabricated magnetic field sensor will be calibrated to see whether it matches the predicted performance. The calibration will be carried out with both DC and AC magnetic modulation. By doing this, both the DC response and dynamic response of the sensor can be tested. The AC test system and signal processing scheme as shown in Figure 2 will be used. The solenoid coils are used to provide a sine wave magnetic field. The sine wave generator provides

the sinusoidal signal to the power amplifier, which converts it to current signal to drive the coils and the lock-in amplifier. The modulated the magnetic field is picked up by the sensor and then demodulated by the lock-in amplifier. By calibration and analysis, the system can be better understood and improved.

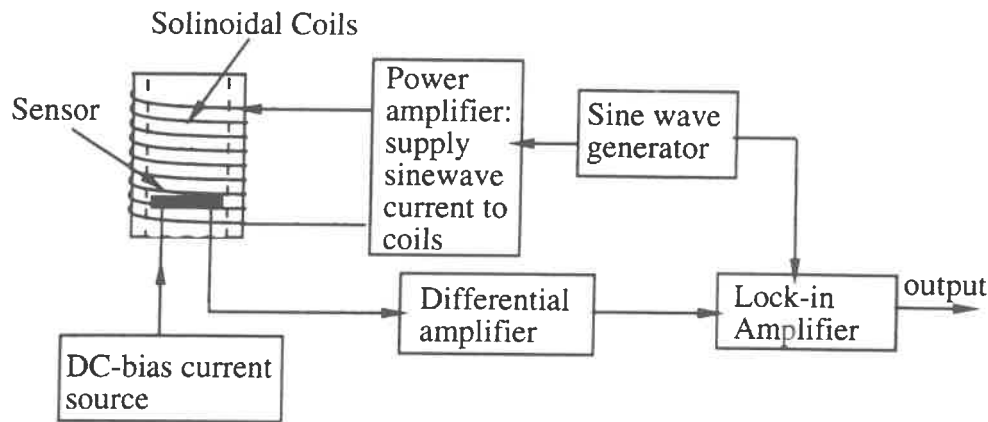


Figure 2. The AC testing scheme for the magnetic microsensor

III. Summary

In summary, we have designed and started the fabrication of a micro magnetic field sensor using the novel GMR multilayer materials and the MEMS technology. We will take advantage of the patterning capability of the lithography processes as well as the inexpensive "thick", electrodeposited GMR multilayers. To the best of the authors' knowledge, this represents the first attempt to fabricate microsensor using GMR multilayer materials. Work along this line is expected to have significant impact on the use of novel magnetoresistive material in the microfabrication field.

Reference

1. S. Kordic, "Integrated Silicon Magnetic field Sensors," *Sensors and Actuators*, 10(1986) pp. 347-378.
- [2]. Zsolt Kadar, Andre Bossche and Jeff Mollinger, "Integrated Resonant Magnetic Field Sensor," *Sensors and Actuators A*, 41-42,(1994) pp. 66-99.
- [3]. S. Kawahito, Y. Sasaki, H. Sato, T. nakamura and Y. Tadokoro, "A fluxgate magnetic field sensor with micro-solenoids and electroplated permalloy cores," *Sensors and Actuators A*, 43 (1994) pp. 128-134.
4. Blondel, A. et al., "Giant magnetoresistance of nanowires of multilayers", *Appl. Phys Lett.* 65 (1994) 3019-3021.
5. Liu, K. et al., "Perpendicular giant magnetoresistance of multilayered Co/Cu nanowires", *Phys. Rev. B* 51 (1995) 7381-7384.
6. Piraux, L. et al., "Giant magnetoresistance in magnetic multilayered nanowires", *Appl. Phys. Lett.* 65 (1994) 2487-2489.
7. Pratt, W.P., et al., "Perpendicular giant magnetoresistances of Ag/Co multilayers", *Phys. Rev. Lett.* 66 (1991) 3060.

Error Analysis by Simulation for a Fiber Based Non-contact Measurement Probe System

Yue Yang, Kazuo Yamazaki

Department of Mechanical and Aeronautical Engineering
University of California
Davis, California 95616

1. Introduction

For an application to non-contact measurement and modeling of a sculptured surface, a fiber-optics based measurement probe which can simultaneously detect surface position coordinates and inclination angles of a sculpture surface was proposed by the authors[1]. The non-contact probe system proposed utilizes LED light emission-reflection principle and assumes that the intensity of the reflection in a certain direction depends on the surface inclination and the distance between the reflection point on the surface and the sensing point of the intensity. Since the probe design integrates the spatial arrangement of one emitting optical fiber with a focusing micro-lens for emitting LED beam onto the surface to be measured and eight optical fibers to receive the reflected light from the surface, the three dimensional alignment of the optical fiber and lens system governs the accuracy of the measurement. Therefore, it is important to understand the relation between the alignment of the fiber system and the measurement characteristics(sensing characteristics of the intensity of the reflected light) of the probe when designing and fabricating the probe system. In order to analyze the effects of the fiber alignment on the detected intensity of the reflected light beam, a simulation model has been developed. Also, experiments have been performed by using the actual prototyped probe to verify the simulation model established.

2. Operation principle and design of the probe system

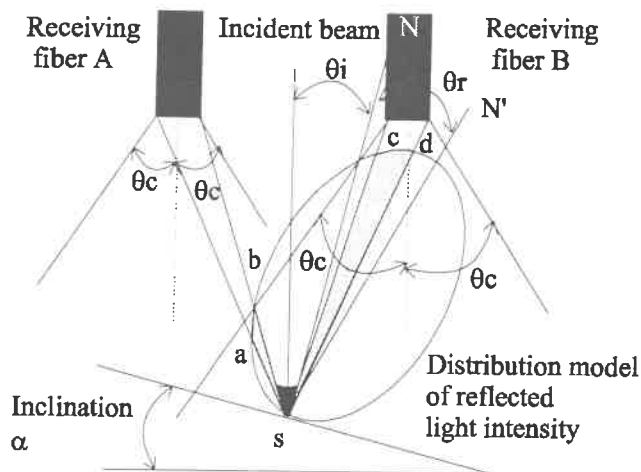
Fig. 1(a) illustrates the principle of measuring a surface inclination. When a incident light beam is applied onto the point S on the surface, the reflected light can be captured by receiving fibers A and B which are symmetrically placed with respect to the incident beam. With the assumption that an intensity distribution of the reflection can be represented by an appropriate ellipse in a two dimensional plane which includes the incident beam and its reflected ray, an amount of the light captured by an optical fiber is proportional to the hatched area(S_{ab} or S_{dc} in Fig. 1(a)). The amount of the light captured by a fiber depends on the number of aperture of a fiber, NA , and only the light which has an angle less than qc with respect to the center axis of a fiber can get into the fiber and the critical angle qc is given by $qc = \sin^{-1}(NA)$ when light enters a fiber from air to a fiber. The relationship between the intensity of the captured light by fibers A and B and inclination angle α can be qualitatively represented by Fig. 1(b). By defining the non-dimensional normalized value Ta as

$$Ta = \frac{Ib - Ia}{Ia + Ib} \quad (1)$$

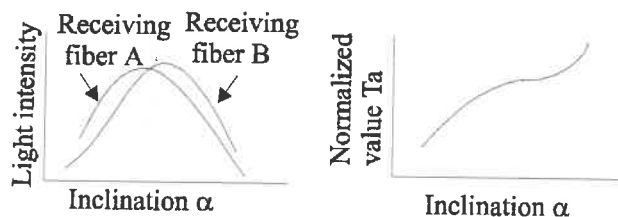
the inclination angle α can be determined for a given light intensity Ia and Ib detected by fibers A and B as shown in Fig. 1(c). The normalized value Ta has a favorable characteristics of being less influenced by a fluctuation of the incident beam intensity and the distance between the fiber and the surface.

For measuring a distance between the probe and the point S on the surface, a method similar to that of the inclination angle measurement can be employed. Fig. 2(a) shows the principle of the measurement. By arranging two optical fibers Ai and Ao which are located on one side of the incident light beam, an amount of light captured by the fibers shows qualitative relationship with the distance Z as shown in Fig. 2(b). In order to obtain a unique relationship between detected intensity and the distance Z , the normalized value Gp is defined as

$$Gp = \frac{Ii - Io}{Io + Io} \quad (2)$$

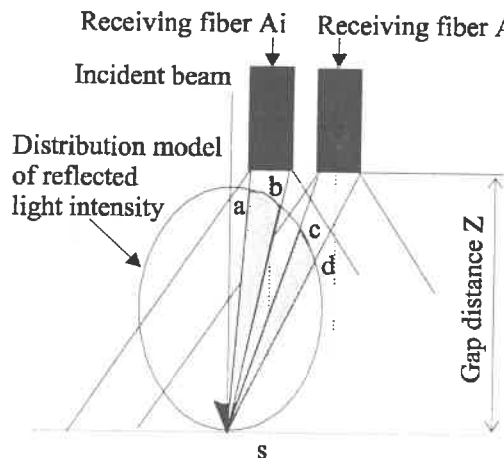


(a) Receiving fibers for detecting inclination α

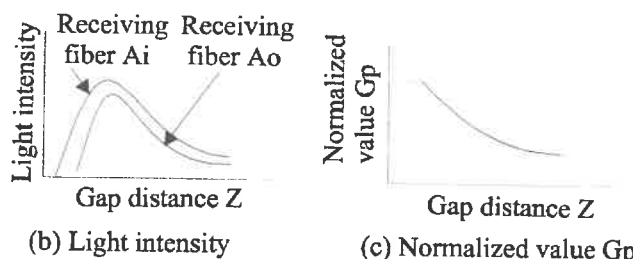


(b) Light intensity (c) Normalized value T_a

Fig. 1 Principle of Measuring Surface Inclination



(a) Receiving fibers for detecting gap distance Z



(b) Light intensity (c) Normalized value G_p

Fig. 2 Principle of Measuring Surface Inclination

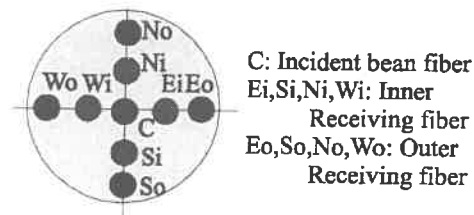


Fig. 3 Fiber Arrangement for 3-D Measurement

where I_i and I_o are the intensity of the detected light by fibers A_i and A_o . A relation between the distance Z and G_p is qualitatively shown in Fig. 2(c). The value G_p has an advantage of being stable against the change of the incident beam intensity and surface inclination angle because of the normalization.

By extending the same principle of two dimensional measurement to three dimensional measurement, the fiber arrangement should be modified such that the two sets of four fibers are orthogonally placed as shown in Fig. 3.

The inclination α of a surface in east-west direction can be detected by fibers E_o , E_i , W_i , and W_o and the inclination ϕ in the direction of north-south can be identified using fibers N_o , N_i , S_i , and S_o . With these two angles α and ϕ , a surface normal vector at a measured point can be determined. The normalized value G_p , T_{ew} , and T_{ns} which are used for determining the gap distance, inclination α and ϕ are defined by equations (3) to (5):

$$G_p = \frac{(E_i + W_i + N_i + S_i) - (E_o + W_o + N_o + S_o)}{(E_i + E_o + W_i + W_o + N_i + N_o + S_i + S_o)} \quad (3)$$

$$T_{ew} = \frac{(E_i + E_o) - (W_i + W_o)}{(E_i + E_o + W_i + W_o)} \quad (4)$$

$$T_{ns} = \frac{(N_i + N_o) - (S_i + S_o)}{(N_i + N_o + S_i + S_o)} \quad (5)$$

where E_i , E_o , W_i , W_o , S_i , S_o , N_i , and N_o are the captured light intensity by respective receiving fibers.

3. The effect of fiber alignment error on intensity detection

The ideal design of the probe needs that all fibers (one emitting fiber and eight receiving fibers) are perfectly aligned in parallel to each other and perpendicular to the end surface of the probe module. When fabricating the probe, however, it is difficult to construct perfectly aligned fibers embedded in the probe. The typical fabrication errors can be classified into perpendicularity error of the emitting fibers to

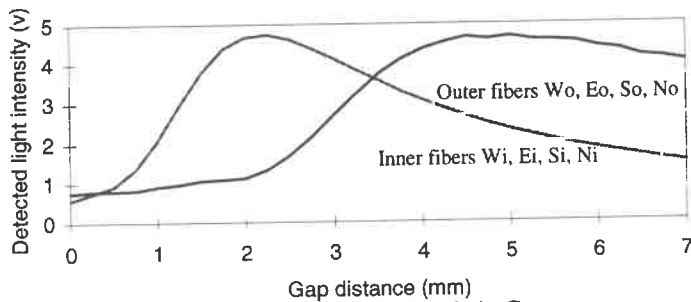


Fig. 4 Idea Characteristic Curve

Ideally, the maximum intensity detected by each of the inner or outer receiving fiber should be reached at the same gap distance z since the fibers are symmetrically distributed with respect to the emitting lens which is located at the center of the module. However, the actual maximum intensity values are not reached at the same gap distance. The difference in gap distances associated with the maximum intensity values of inner fibers or outer fibers is referred to

as the *characteristic difference*, which has a direct influence on the measurement accuracy[2]. The cause of the characteristic difference can be considered as fabrication errors.

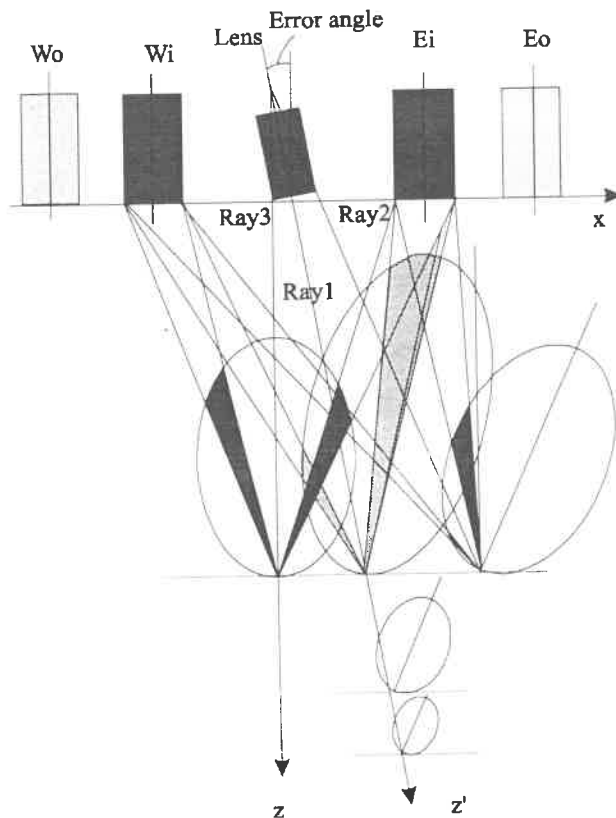


Fig. 5 Geometric Illustration of the Simulation Model

associated with the center ray are illustrated in the figure.

The simulation of light intensity detection is based on the calculation of the detectable area of the reflected light intensity distribution. To determine the detectable area, we must consider the shaded areas in Fig. 6 which are the intersection areas of the ellipse and the straight line segments from the receiving fibers. If a shaded area is within the critical angle of the receiving fiber, the area is determined to be the detectable area. The detectable area can be calculated using the following formula:

the end surface and position error of the receiving fibers with respect to the center of the emitting fiber.

The influence of the position errors of the receiving fibers is negligible because the value of the position error is small compared to the gap distance z . Therefore, it may be deduced that the characteristic difference in the characteristic curve plot is mainly caused by the perpendicularity error of the emitting lens.

Based on the above analysis, a three dimensional simulation model was developed. Fig. 5 is a two dimensional geometric illustration of the simulation model. Every ray from the lens is supposed to produce an ellipse. Actually, there are infinite rays distributed within the extreme boundary rays which can be calculated by using matrix equation for ray tracing in the lens[3]. Finite rays can be used to approximate infinite rays provided that proper angle interval is chosen. Only the center ray and two extreme rays in the right-most and left-most are drawn in the figure. For clarity, only the detectable areas associated with the two inner fibers are marked. Besides, the sizes of the ellipses are different for different gap distance z and only some of the ellipses

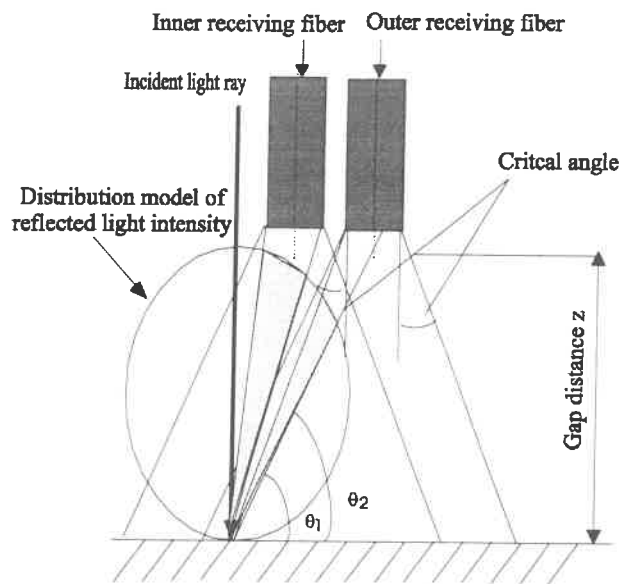


Fig. 6 Simulation of Light Intensity Detection

In order to verify the applicability of the simulation, the experimental study has been conducted by fabricating the probe. The geometrical dimensions of the probe fabricated is shown in Fig. 7. In the prototyped probe, there are five modules. The one at the bottom was used for the study. The experiments were conducted by attaching the probe to a three dimensional coordinate measuring machine. In the experiments, the actual characteristic curves were obtained by applying the probe to an appropriate surface specimen. Using an appropriate reflection distribution model which is close to that of the actual specimen, the simulations were performed for different module parameters.

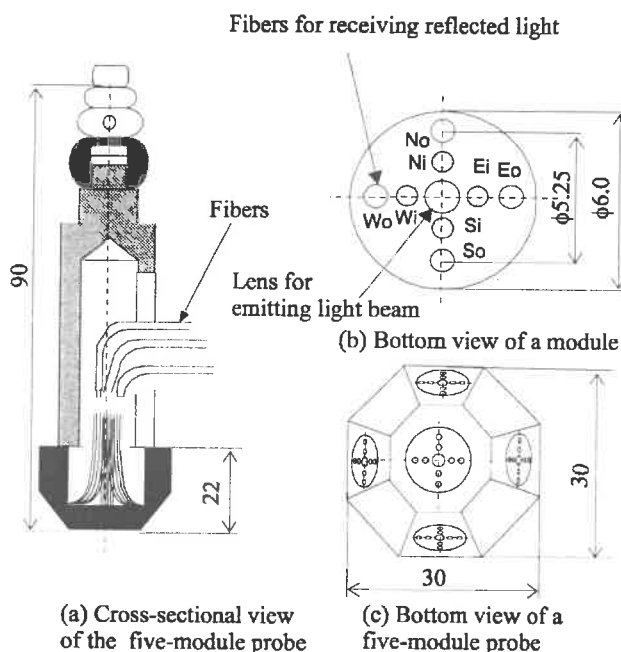


Fig. 7 Probe and Module Structure

$$\int_{\theta_1}^{\theta_2} \frac{2a^2b^4 \sin^2(\theta)}{(a^2 \cos^2(\theta) + b^2 \sin^2(\theta))^2} d\theta \quad (6)$$

where a and b are the semi-axes of the ellipse of the reflected light intensity distribution and can be chosen according to the reflection characteristics of the surface to be measured. The integral (6) can be calculated using Equation 7:

$$A = b^2 \left[\frac{a}{b} \tan^{-1} \left(\frac{\tan(\theta)}{\frac{a}{b}} \right) - \frac{\left(\frac{a}{b} \right)^2 \tan(\theta)}{\tan^2(\theta) + \left(\frac{a}{b} \right)^2} \right]_{\theta=\theta_1}^{\theta=\theta_2} \quad (7)$$

4. Application of simulation to the study on the characteristics of the prototyped probe

Fig. 8 shows the comparison of the detected light intensity values of the actual probe to those of the simulation. They have a good agreement with each other within the dynamic range of measurement. Therefore, the calculation of detectable area provides a valid simulation of light intensity detection using the developed fiber-optic based non-contact measurement probe.

A quantitative relationship between the error angle and the characteristic difference was obtained using the proposed simulation method. Fig. 9 is the result for an error angle range from 0 to 15 degrees. The relationship provides a critical reference of actual module parameters. For example, the actual characteristic differences of the module used in the experiments are about 0.75 mm for the east and west fibers and 0.6 mm for the north and south fibers respectively as shown in Fig. 10, which, by referring to Fig. 9, implies that the error angle of the emitting lens is about 8.5 degrees in east-west direction and 7 degrees in north-south direction respectively.

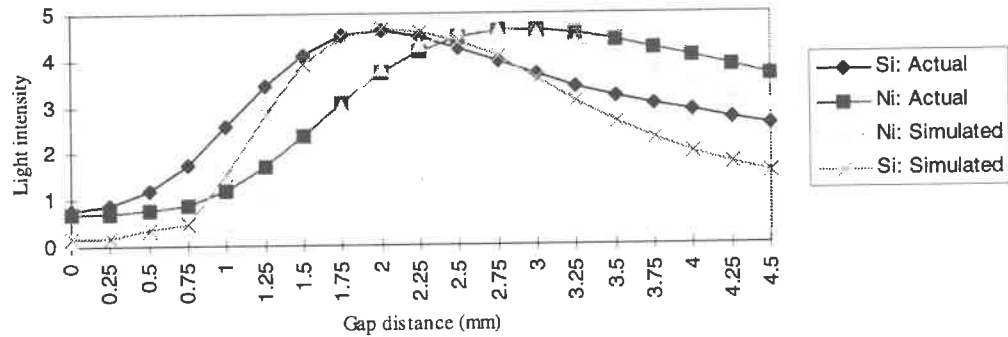


Fig. 8 Comparison of actual and simulated light intensity values

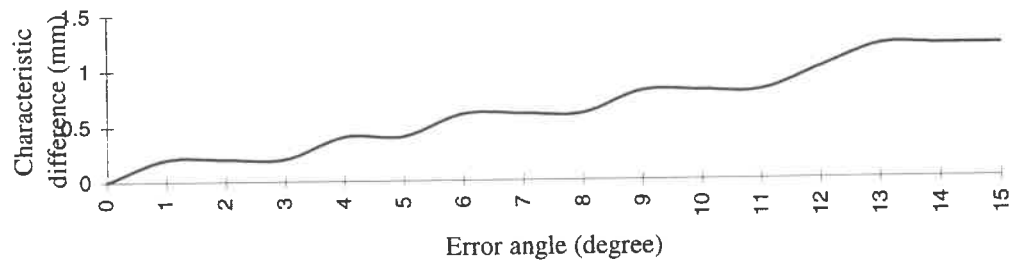


Fig. 9 Relation ship between Error angle and Characteristic Difference

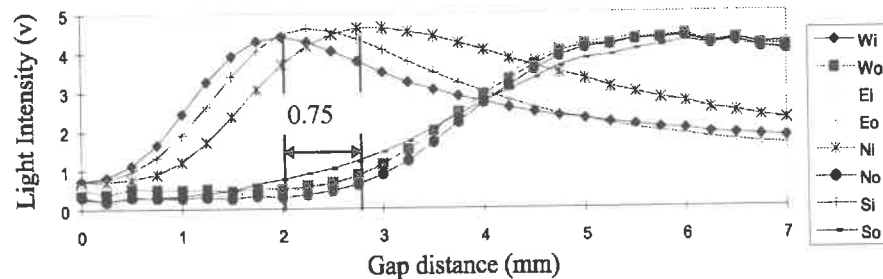


Fig. 10 Characteristic Curve Plot

The characteristic differences have an important influence on the measurement accuracy of the developed probe system. For example, the characteristic differences in Fig. 10 will produce a non-zero surface inclination angle when measuring a horizontally-positioned flat surface and this error. The availability of the relationship between the module parameters and the characteristic differences may provide a valid evaluation of the fabrication of the fiber-optic based non-contact measurement probe.

4. Conclusions

A study on an effective error analysis method for the optic-fiber based non-contact measurement probe has been performed. The followings are the conclusions of the study:

1. A simulation model of the probe was developed based on the assumption of elliptic distribution of reflected light intensity.
2. The quantitative relationship between the error angle of the emitting lens and the characteristic difference was obtained by the developed simulation methods.

3. The simulation results can be used to evaluate the design parameters of the probe.
4. The simulation result supports the assumption that the characteristic differences of the actual probe were caused by the fabrication error of the probe.
5. The developed method is useful to enhance fabrication accuracy of the probe by using the simulation results.

Reference:

1. Kazuo. Yamazaki et al., "Noncontact Probe for Continuous Measurement of Surface Inclination and Position Using Dynamic Irradiation of Light Beam", The Annals of the CIRP, Vol. 42/1/1993
2. Yue Yang et al, "Non-Contact Autonomous Measurement System", Technique Report, Intelligent Manufacturing System Laboratory, University of California, Davis, 1995
3. Kee Sein Lee, "Development of Non-Contact Probe for Continuous Measurement of Surface Inclination and Position", Thesis of Master's Degree, University of California, Davis, 1992

Micro-volume Defects Measuring System Based on Digital Image Processing Measuring Technique

You Zheng Yang Ren Chen Jun

Department of Precision Instrument Tsinghua University, 100084, Beijing, P.R.China

ABSTRACT

Scattered light by micrometer size particles has some special characters. This paper provides a volume defects measuring technique in silicon materials using Mie theory. According to the infrared transmission property in silicon materials, the infrared laser and infrared CCD camera are used. An arrangement of the measuring device is given.

Key words : Mie theory, infrared, volume defects, scanning tomography, contrast

1. INTRODUCTION

With the development of semiconductor electronics, micro-machine and micro-system, semiconductor and its mechanical properties are required a still higher quality. So their corresponding measuring technique is required further perfected. Now although there are many measuring methods for surface roughness or surface micro defect, such as STM, AFM, optical and mechanical profilometer, for micro defects inside materials the usual measuring method is also transmission electron microscope. Electron microscope is suitable for the samples with high density of defects, and at the same time, it is complex and destructive for the sample. So in this paper a novel micro-volume defects measuring technique in silicon materials using Mie theory is proposed.

2. Defect and Infrared property of silicon-based semiconductor crystal

Wafers of silicon or germanium must have some defects during the processes such as single-crystal growth and so on as listed in Tab.1

Tab. 1

Process	The main induced defects	The effect caused by induced defects
cutting into slices	tool mark, injury, micro-crack	wrapping, slice-smashing, contamination
grinding	fringe-smashing, micro-crack	slice-smashing, contamination, bulk injury
polishing	scratching, residual stress	bulk injury, surface injury
extension	fault, dislocation, micro-defect, gliding	strain, wrapping, pernicious occluded sediment
oxidize	fault, dislocation, fog defect	wrapping, occluded sediment
diffuse and heat treatment	emitter fringing dislocation, mismatch dislocation	straining, wrapping, emitter dipping effect
ion implantation	secondary defect	impurity, sediment, injury, wrapping
CVD	secondary defect	strain, wrapping
metallization		strain, contamination, migrancy

interlinking	secondary defect	strain , crack
encapsulation		injury , contamination

Semiconductor crystals such as Ge 、 Si 、 CdTe 、 GaAs and so on are transparent for infrared light . A typical Si transparent property curve is given in Fig.1. From the curve , it can be seen , when the wave length is from 1.2 μm to 1.6 μm , the infrared light can go through the silicon materials very easily.

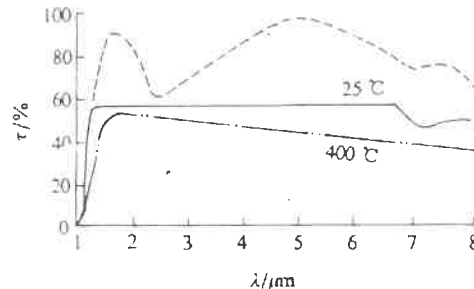


Fig.1 The infrared transparent property of silicon

3. Mie theory and the new method of measuring micro-volume defects in the silicon material

The new method of measuring micro-volume defects in silicon materials is given based on the foundation and development of Mie scattering theory .

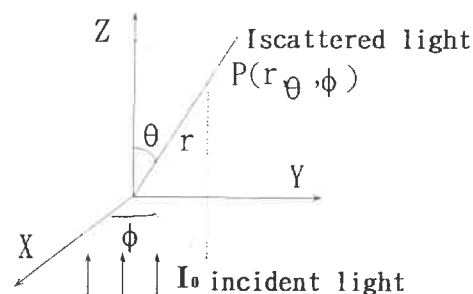


Fig.2 The coordination system of the spherical particle

A single spherical particle is placed at the zero of a rectangular coordination system (In Fig.2). The particle is illuminated by a plane wave . The incident light is parallel to Z axle . The scattered field is determined by the size parameter $q=2 \pi R / \lambda$, where R and λ represent the radius of particle and the wavelength in the surrounding medium of the particle and by $m=n-j \eta$, the ratio of the complex refractive indices of particle and surrounding medium . The observing plane is formed by the direction of the incident light and of the scattering light , we can get the equation which represents the intensity at $P(r, \theta, \phi)$:

$$I(\theta, \phi) = \frac{I_0 F(\theta, \phi, q)}{2K^2 r^2} \quad (1)$$

where I_0 and θ represent the intensity of incident light and the scattered angle , $K=2 \pi / \lambda$ and r represents the distance between the single particle and the point P . $F(\theta, \phi, q)$ is called scattering function , which can be showed as :

$$F(\theta, \phi, q) = i_1(q, \theta) \sin^2 \phi + i_2(q, \theta) \cos^2 \phi \quad (2)$$

where i_1 and i_2 are called Mie intensity index which are the functions of the scattered angle θ , the complex refractive index m and the size parameter q .

From (1)(2) , it can be seen that the light intensity in scattered field only rely on q , m , θ . If q , m are constant , the scattered light intensity is related only with the scattered angle θ , it can be said that the intensity will chang only when the receiying direction changes.

The reliable theoretical basis is given by Mie theory for measuring defects in materials used scattered light technique . The micro defects in materials , for example , mental polymerization and micro desposit become the scattering center , and the scattered light will be caused by them . So , if we recorded the scattered light intensity in arbitrary direction and solid-angle , we can get the distribution of defects .

This idea was tested by R.M.Sliva in 1984 . Sliva checked the surface defects and the near surface defects of a GaAs diaphragm with 2 inch diameter by using the photon back scattering technique. He got the distribution of the defects . In Fig.3 , the Sliva photon back scattering system principle is given .

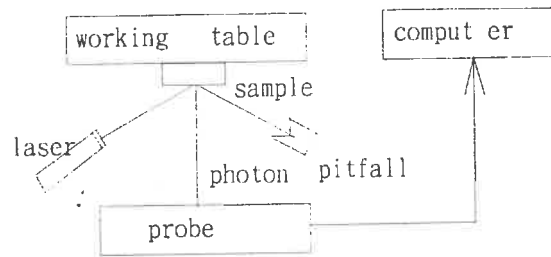


Fig. 3 The principle of the photon back scattering system

With the development of digital image processing technique , it is possible that the information of smaller size volume defect can be gotten by the image of an optical microscopy .

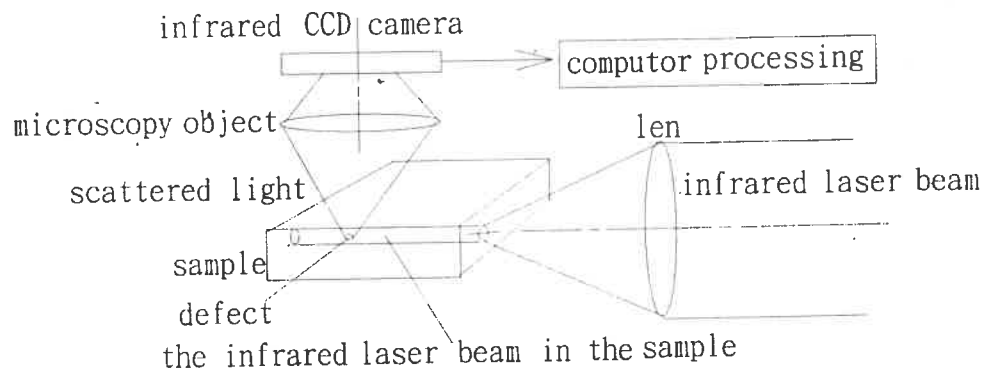


Fig.4 The principle of the infrared laser scanning tomography

According to Mie theory and with the help of infrared laser technique and digital image processing technique , we developed a new micro-volume defects measuring technique in silicon material using photon scattering and infrared property of semiconductor . It is called infrared laser scanning tomography . The principle of this method is that a measured section of the sample was scanned by infrared laser beam , with the infrared light that is transparent for semiconductor is scattered by volume defects in the sample , the scattered light is gathered at the direction that is orthogonal to the section . As this , a defect diagram can be gotten . Changed the scanning depth of laser in the sample , a tomography diagram can be gotten which represents the distribution of the defects , and the density and location of the defects can be gotten by digital processor . These changed diagrams are digitally stored with 512 pix per line and 256 grey levels , then a two dimensions diagram

(512 × 512pix) which represent the defects' distribution is formed. In Fig.4, the principle of infrared laser scanning tomography system is given.

The main characteristic of this technique is its high contrast. If I and I_0 respectively represent the grey of image and background, the contrast $C=(I-I_0)/I \times 100\%$ of this new method in detecting defects in material will be 100% even when defects are smaller than 1 μm because of the high intensity of infrared laser and the low background as scattered light is collected vertical to the incident. So when there are defects in the sample, according to Mie theory, the defects' image can be gotten. This is illustrated in Fig.5. On the basis of Fig.2 and equation (1), (2), at O point, $\theta=90^\circ$, $\phi=90^\circ$, the light intensity only related with m and q , it can be said that the different scattered light intensity for the different character and different size of the defects. From Mie theory, we can get that for different character and different size of the defects, although r is the same, the scattered image will be different. Because the location of the defects can be determined by a precise three-dimension stage, the size parameter q can be easily got according to equation (1).

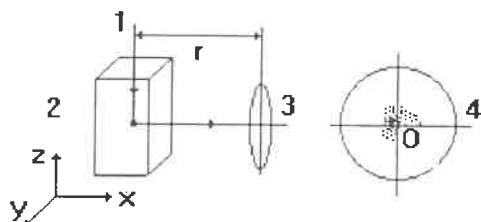


Fig.5 The scattering image of the micro defect

4. CONCLUSIONS

In a word, researching the infrared technique of measuring micro-volume defects in semiconductor materials has very high practical value. At first, it is not only applied to the instruct analysis from the interior of the material, but also can get the density and location of the defects. So, when it is used for analyze the growth parameters and the relation between the processing difference and producing defects after growth, the theatrical base of semiconductor defects electronics and micro machine design can be provided. Secondly, if this technique develops perfectly, it probably became the technique on line of measuring the micro-volume defects in semiconductor materials.

REFERENCE

1. Junxin Fang and Lu Dong, Solid Physics, Shanghai Science Publication House, December, 1980
2. G.Mie, Annalen der physik, 1908, Vol.25, No.25, 377-378
3. R.M.Sliva, A new instrument for constant $(\beta - \beta_0)$ scatter mapping of contiguous optical surfaces of up to 25 square inches, SPIE, 1984, 511, P38-43
4. Jitao Wang and Mingdeng Liu, Semiconductor materials, Higher Education Publication House, May, 1990
5. Shuquan Wang, Infrared Measuring Basic, Shanghai Communication Publication House, March, 1991

A Study on High-Accuracy Laser Interferometric Measurement of Micro-Angle

Zhu Hongxi Wang Haibo Liu Hongxiang
The Institute of Precision Instruments and Electro-Optics Engineering,
Tianjin University Tianjin30072 China
Tel: 0086 22 7487459 Fax: 0086 22 3358329

Abstract

In this paper, the concrete method of the laser interferometric measurement of micro-angle is presented, and a new type of double reference mirrors laser interferometric measurement apparatus is advanced, the angle measurement principle and the distinguishing feature as well as the concrete application of this apparatus are discussed in detail.

Key words: laser interference measurement reference mirror

1. The basic characteristic of micro-angle measurement

The general principle of micro-angle measuring apparatus is showed in Fig.1, the measuring equation can be expressed as:

$$\operatorname{tg} \alpha = \frac{h}{l} \quad \text{or} \quad \sin \alpha = \frac{h}{L} \quad (1)$$

as the angle is so small, then $\alpha = \frac{h}{l}$ or $\alpha = \frac{h}{L}$ (2)

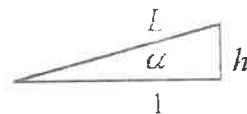


Fig.1

obviously, the measurement of micro-angle α is to measure h under the determined l (or L) or simultaneously measure h and l (or L) with an adequate accuracy measurement way.

By equation (2), we get the measurement error equation of the α to be measured:

$$\Delta \alpha = \frac{1}{l^2} (l \Delta h - h \Delta l) \quad (3)$$

where l is determined by the structure of the measurement apparatus or the size of viewing field. Under given measurement accuracy, the bigger the l is, the higher the

measurement accuracy of measured α is. The micro-angle generator is just designed based on this principle.

Considering the measurement accuracy matching and the optimal measurement condition, we ought to take $l\Delta h = h\Delta l$ (4)

for the micro-angle measurement, $l \gg h$, then $\Delta l \gg \Delta h$. therefore, in the measurement of micro-angle, it required more higher measurement accuracy for h than l , the smaller the measured angle is, the bigger the difference between the required measurement accuracy on l and h is. e.g., when the angle to be measured is $2''$, $\Delta l = 1724\Delta h$.

Obviously, the key to upgrade the micro-angle measurement accuracy is to upgrade the measurement accuracy of h . so it is always a research direction in the micro-angle measurement to which researchers have been paying a good deal of attention.

2. The principle of micro-angle laser interferometric measurement

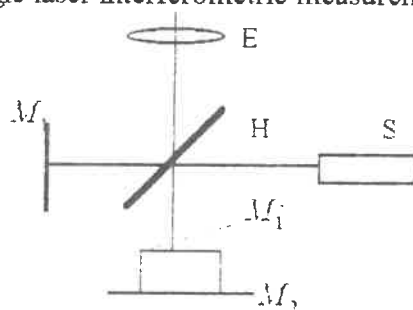


Fig.2

As showed in Fig.2, the laser beam, which diameter is expanded to $\phi 50mm$, projects on the beam-splitter mirror H and be splited apart into 2 beams, one is the transmitted beam and the other is the reflected beam. the transmitted beam is reflected back along the original path on the reference mirror M_1 , while the reflected beam turns 90° and projects on the surface of the optical flat M_2 which set on the work table and be reflected back, so the two beams interfere when they meet again. By the eye lens E , we can see the parallel interference fringes with equal intervals. As the hight of the air wedge between two adjacent fringes is $\lambda/2$, so the hight of the air wedge corresponding to n fringes is $n\lambda/2$. therefore, the angle included between M_1' and M_2 is :

$$\alpha = \frac{h}{l} = \frac{n\lambda}{l/2} \quad (5)$$

where M_1' is the image of M_1 in the beam-splitter mirror H and l is the intervals width corresponding to n fringes.

Now we set an angle gauge with the angle β on the surface of the optical flat M_2 as showed in Fig.3, where M and M_2 are the work surfaces of the angle to be measured, M_1' is the image of the reference mirror, the angle included between M_1' and M is α_1 and the angle included between M_2 and M is α_2 , thus we can see two sets of interference

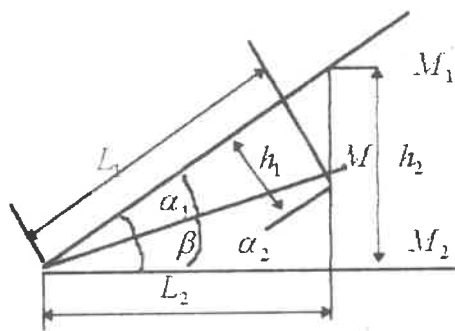


Fig.3(a)

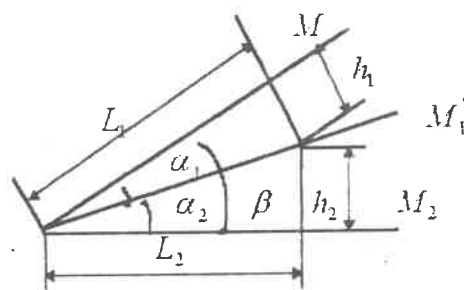


Fig.3(b)

fringes as showed in Fig.4 by the eye lens.

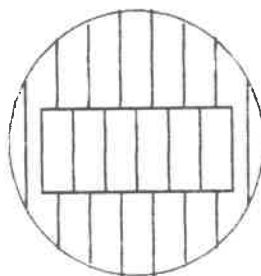


Fig.4

$$\alpha_1 = \frac{n_1 \lambda}{l_1 2} \quad (6)$$

$$\alpha_2 = \frac{n_2 \lambda}{l_2 2} \quad (7)$$

where l_1 and l_2 are the interval width corresponding to n_1 fringes and n_2 fringes respectively, we take $l_1 = l_2 = l$ as the selected length on the work surface of the angle to be measured, n_1 is the fringes number in the length l on the work surface of the angle to be measured and n_2 is the fringes number in the length l on the surface of the optical flat, then we get

$$\text{for Fig.3(a)} \quad \beta_1 = \frac{\lambda}{2l} (n_2 - n_1) \quad (8)$$

$$\text{for Fig.3(b)} \quad \beta_2 = \frac{\lambda}{2l} (n_2 + n_1) \quad (9)$$

gently press the eye lens tube, we can see these fringes moving. When the fringes move in the same direction, we use equation (8) to calculate the measured angle. If the fringes move in the opposite direction, we use equation (9) to calculate the measured angle. The theoretical study and experiment show that the measurement error is smaller than $\pm 0.3''$ when the angle to be measured is not bigger than $2''$.

3. The principle of high accuracy measurement of the micro-angle using double reference mirrors

As showed in Fig.5, we substitute the reference mirror shown in Fig.2 for a couple of juxtaposed reference mirrors and adjust tilt angle of the two reference mirrors to make the included angles between their images and the work surface of optical flat are all β , we get two sets of equinterval interference fringes. If the optical flat turns an angle α to its original position, the angle included between the reference mirror M_1 and the optical flat is $(\beta - \alpha)$, while the angle included between the reference mirror M_2 and the measuring mirror is $(\beta + \alpha)$, the interval widths of the two sets of fringes are

$$e = \frac{1}{2n(\beta - \alpha)} \lambda \quad (10)$$

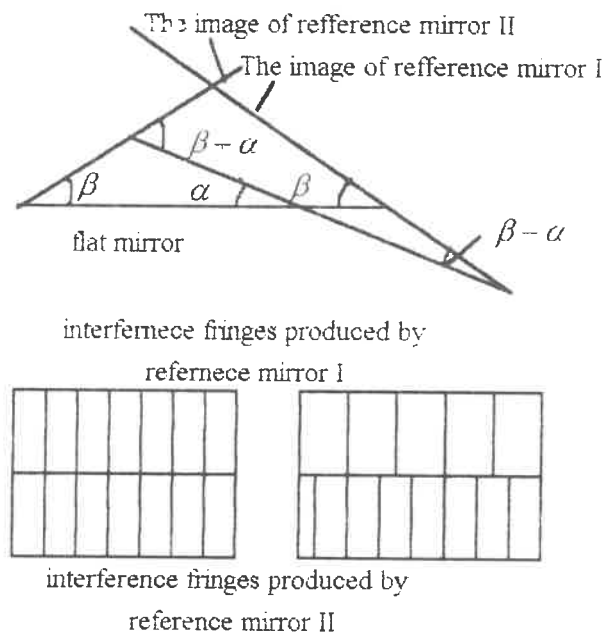


Fig.5

$$e' = \frac{1}{2n(\beta + \alpha)} \lambda \quad (11)$$

respectively. The interval widths of the two sets of fringes change with the characteristic of differential.

Supposing

$$e = ke'$$

i.e., $k = \frac{e}{e'}$, apply it to equation (10) and (11), we get

$$\alpha = \frac{k-1}{k+1} \beta \quad (12)$$

from the angle β and the changes of the fringes interval, the angle α can be determined. Usually β is very small, it is determinate after the reference mirrors is adjusted properly and can be accurately measured with the method mentioned above, i.e., from the interference fringes number n in a certain length l , it can be calculated with the under expression

$$\beta = \frac{n}{2l} \lambda \quad (\text{rad}) \quad (13)$$

generally, 1/10 width of the fringes interval can be resolved by eyes.

$$\begin{array}{lll} \text{if} & \beta = 1'' & k=1.1, \quad \alpha = 3'' \\ & \beta = 30'' & k=1.1, \quad \alpha = 1.4'' \\ & \beta = 20'' & k=1.1, \quad \alpha = 1'' \\ & \beta = 2'' & k=1.1, \quad \alpha = 0.1'' \end{array}$$

Obviously, the magnitude of β prescribes the measurement scale, the smaller the measurement scale is, the higher the resolution is. For the distinguishing feature of the double reference mirrors interferometric measuring apparatus, we can use it in the measurement of the flatness and straightness of precision slides and the polyhedral prism calibration as well as precision positioning, especially for the polyhedral prism calibration with the method of the whole combination-permutation and mutual comparison, we can even use it to measure out the tower error of polyhedral prism if necessary, so the tower error of polyhedral prism has no effect on the positioning accuracy. This is a distinguishing advantage of the double reference mirrors method.

More over, as we get the measurement result in this method by comparing the widths of two sets of fringes formed under the same measurement and environment condition, so the environment and temperature as well as atmospheric pressure have little effect on our measurement result.

4. Conclusion

When the measurement scale is not too much, the method of laser interferometric measurement of micro-angle can reach a rather high measurement accuracy, usually it is only used in micro-angle mutual contrast, its utility is limited.

The method of di-reference mirrors measurement of micro-angle posses a strong anti-interference feature, when the measurement scale is small, it has a very high resolution power. For high accuracy collimate position and flatness as well as straightness measurement, it opens up a new way with wide foreground in exploitation and applying.

STYLUS DRAG FORCES ON LUBRICATED SURFACES

D. G. Chetwynd and L. A. J. Davis

Centre for Nanotechnology and Microengineering
Department of Engineering,
University of Warwick, Coventry CV4 7AL UK

Introduction. Despite increased competition from non-contacting methods, stylus-based surface profilometry is still the preferred method for many applications. Thus there remains interest in improving its traditionally perceived weaknesses: speed of measurement and surface damage. At higher traverse speeds there is a tendency to lose fidelity of profiling (the stylus may leave contact at the tops of asperities) and to scratch the surface (through higher dynamic forces or if the static contact force is increased to reduce stylus bounce). There is evidence [1] that quite high damping factors in the pick-up are helpful. Stylus friction is relevant to this as it provides one energy dissipation mechanism. However, friction is likely to have greater influence on fidelity through stick-slip, etc. and on surface damage since the classic Hertzian stress field is modified by the frictional force, bringing peak stresses towards the surface at the trailing edge.

Recent study of stylus behaviour for general instrument performance has shown that the measured lateral force (taking the profile signal to be in the vertical plane) is not a direct measure of the stylus frictional drag but a complex signal of this and geometric effects [2]. Friction coefficients behave largely as expected of macroscopic contacts. However, the published results generally arise from surfaces that have been cleaned at least by wiping and so are free from gross lubrication, although there may be adsorbed films. There is anecdotal evidence that in practice the stylus region is sometimes flooded with mineral oil and that, for example, glass components are sometimes 'cleaned' by wiping with the hand, which will transfer a trace of natural oil to them. There appears to be virtually no discussion of using deliberately applied lubrication, which may indicate either that it has not been examined systematically or that only inconclusive results were found. This is hardly satisfactory and so we report here on the effects of various forms of lubrication when measuring the roughness of reasonably smooth engineering surfaces.

Lubricated Stylus. Taking the contact as a Hertzian one of sphere on flat, a 2 μm radius diamond stylus imposes direct stresses often at the level of 10 to 20 GPa, well above the yield point of bulk metals. By conventional guidelines [e.g. 3], boundary lubrication rather than elastohydrodynamic will apply even if there is free oil around the stylus. Intuitively, a trace of oil might reduce surface damage under higher traverse speeds because of reduced friction. With a flood of oil conditions below the tip will be little changed and lower friction maintained but the total lateral force might rise through viscous drag on the stylus body. Hydrodynamic effects are unlikely at the range of speeds and forces prevailing here. Oil may also help maintain fidelity of measurement through:

- lower stick-slip and so less high frequency noise;
- stylus motion in the oil adding damping to the system;
- surface tension effects reducing lift-off at peaks (non-linear damping effect).

Potential disadvantages of oil layers include the need to maintain cleanliness for some workpieces and reduced fidelity from direct filling or hydrodynamic skating over small valleys.

Experimental Method. Stylus drag forces were measured with the same system as used for earlier studies of stylus dynamics [2], figure 1. A Rank Taylor Hobson Talysurf 5 profilometer was modified by attaching a small moving-magnet electromagnetic force actuator to its stylus arm. This allows the contact force to be varied with only a slightly reduced dynamic response (lowered resonance of the pick-up) and minimal noise injection since the compliant actuation effectively decouples vibration. Small specimens are mounted on a monolithic notch-hinge elastic mechanism, made of aluminium. Stylus drag slightly deflects its stage, measured by a nanometre sensitivity inductive gauge, based on a Talystep sensor head. Working at very high sensitivity exacerbates noise problems but offers the advantage that, in the normal force regimes being studied, the horizontal deflection is small compared to the nominal stylus tip-size and so profiles from the Talysurf are not noticeably distorted laterally by the force detection. Data acquisition was by a 12-bit analog to digital converter into a 486-based personal computer, running the LABVIEW package. The normal profile and the output of the force sensor were sampled at constant time intervals, assuming constant traverse speed conditions from the Talysurf 5, corresponding to intervals of $0.5\ \mu\text{m}$ on the surface, unless otherwise stated.

The force actuator and the lateral force cell were calibrated against each other since this study is largely concerned with coefficient of friction for which their ratio is required. The flexure was placed with its axis vertical and small laboratory weights hung from it on cotton thread in order to obtain its basic calibration to within 5%. Then, in the same orientation, the stylus was rested against it and the current in the actuator coil calibrated against its output. Forces may be set reliably to within $10\ \mu\text{N}$. The resolution of the digitized drag signal is well below $1\ \mu\text{N}$ but instrument noise limits useful resolution to around $2\ \mu\text{N}$.

Specimens were chosen to represent a range of practical measurement situations while maintaining conditions for which friction effects are not unduly masked either by resolved

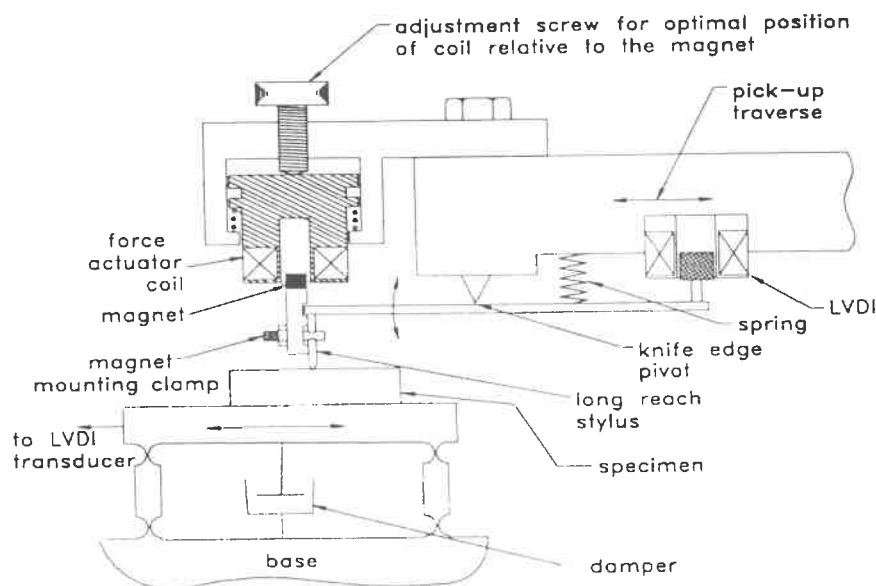


Figure 1: Schematic of the stylus-drag measurement system.

components of the normal force deriving from local geometry or by instrument noise. Mild steel and aluminium specimens were first ground and polished and then slightly roughened by a fine-grit abrasive paper to give an R_q value of around $0.2\text{ }\mu\text{m}$ at 0.8 mm cut-off. These contain short-wavelength features that help to test profile fidelity and the softer aluminium is more vulnerable to stylus-induced damage. Smoother surfaces would better show the direct effects of lubrication upon the friction coefficient and any lateral viscous drag and here glass microscope slides and a steel gauge block were used to give variations in both hardness and surface chemistry.

Lateral force was measured along with the profile so that direct comparisons could be made, but the friction force was estimated by making each trace a combination of a forward and reverse traverse. The force trace may contain a linear trend from imperfect alignment of the specimen to the traverse axis and this was removed by fitting a linear least squares straight line separately for each direction. Then the mean difference of the remaining signals was taken as the friction force since friction effects (always opposing motion) reverse with the traverse but geometrical ones do not.

The lubrication conditions were set up by first cleaning each specimen in a standard way by flooding with acetone, wiping with lint-free cloth and then allowing a brief period for evaporation of any remaining solvent. Trace lubrication was then applied by flooding the surface with the appropriate oil and wiping off with tissue. Flood lubrication consisted of lifting the stylus clear, adding a blob of oil and then lowering the stylus gently by means of the force actuator. Initial contact with the oil was detected by the change of the force-displacement sensitivity in the approach and film thickness could be estimated from the further drop of the stylus from this point until the working load was established. This method does not account for the meniscus that forms around the stylus shank and so a second estimate of film thickness was obtained by using a low-power, long-focus microscope. Both methods are prone to error, but they are adequate for checking whether the behaviour of the oil follows reasonably that expected of a Newtonian fluid.

Results. The profiles obtained from the roughened surfaces on steel and aluminium with different forms of lubrication but otherwise under nominally identical conditions, typical of those used in normal practice, were generally indistinguishable at the scale that could be reproduced here and examples are not shown. On detailed examination, minor divergences that might be attributable to lubrication conditions rarely exceeded 2% of amplitude on these profiles (where peak to valley amplitudes are around $1.5\text{ }\mu\text{m}$). Slight surface damage was seen on aluminium measured at a speed of 1 mm s^{-1} , rather more so on clean surfaces than lubricated ones. Quantification of damage is generally problematic and not attempted for the small effect seen.

On no specimen was the friction force distinctly influenced by any experimental variable (except stylus force). Neither drag forces nor friction coefficients showed statistically significant variation with lubrication conditions. Cluster analyses show a slight tendency to group according to lubricant, but there are many 'outliers' from the main groups: large data sets will be needed to draw firm conclusions and then time dependent behaviour may confuse the issue. Figure 2 shows a typical pattern (and one of the clearer trends found). With a 1 mN load on a steel gauge block, clean conditions gave the lowest forces (around $60\text{--}80\text{ }\mu\text{N}$), a trace of oil increased this a little and flooding with oil to a depth of around 0.5 mm gave a further increase (around $100\text{--}110\text{ }\mu\text{N}$). Increasing oil depth to around 1 mm reduced drag slightly. This, and the lack of correlation of drag with speed, suggests that there is very little viscous drag on the stylus. Setting

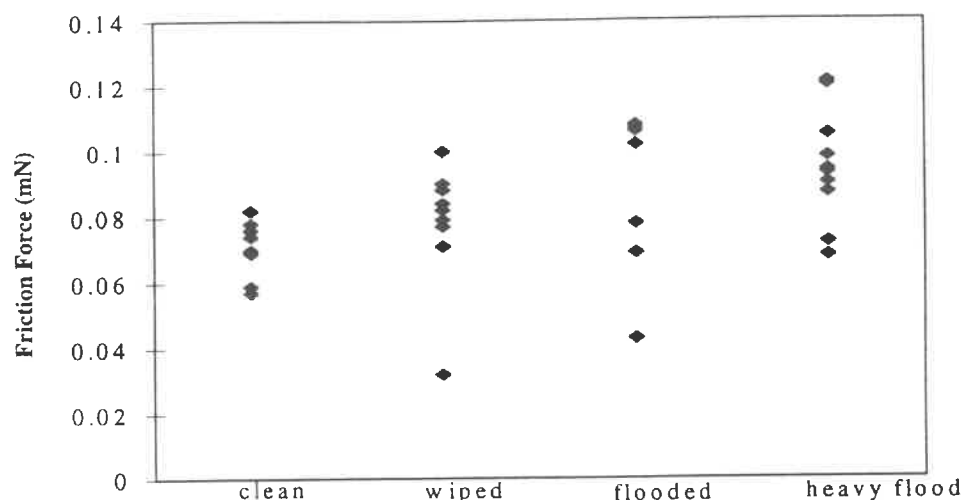


Figure 2: Friction versus lubrication conditions: 1 mN stylus load on a smooth steel surface.

the force actuator to the minimum not causing gross lifting of the stylus arm should result in drag representing only viscous effects and around 10 μN was found with 300 cP oil, but with considerable scatter this near to the resolution limit. On glass the trend was even less clear: clean and trace-lubricated surfaces showed no systematic grouping. Although having slightly higher mean value, the flooded ones clustered towards the middle of the complete set. With a 5 mN stylus load, the flooded cluster was rather more distinctly above the others, which remained indistinguishable. The friction coefficient at the higher load (about 0.04) was roughly half that at 1 mN. The aluminium specimens, measured at 0.5 mN stylus force, gave a coefficient of friction of around 0.25 with a slight tendency for the clean surface to cluster at the higher values (120–140 μN) and the flooded ones towards the lower values (100–120 μN), while the trace lubricated ones were scattered through the range. The viscosity of the oil, at least in the range 50 to 300 cP, appears unimportant.

Even if these trends are real, the effect of lubrication is not of much practical importance for the normal operation of stylus profilometers. The results are consistent with the pressures at the tip being so severe that almost all lubricant is excluded, so that there is little difference at the interface however much oil is in its vicinity. Where lubricated surfaces show marginally higher friction, highly non-Newtonian behaviour of the oil at extreme pressures may be increasing the real area of contact without reducing the effective shear strength. However it might merely reflect the practical difficulty of properly cleaning the stylus while it is on the instrument: remnant traces of lubricant could suffice to make the tip conditions similar to the fully lubricated ones if virtually no oil gets to the interface under any of the conditions studied.

Using flood lubrication does affect the initial behaviour of the instrument, figure 3. The stylus settles through the oil level to detectable limits in no more than five seconds. However taking a first trace about 10 s after lowering the stylus gives an inaccurate profile. Reversing and repeating without lifting the stylus then gives profiles corresponding well to each other and to those taken without lubrication. Lifting the stylus briefly re-introduces an inaccurate form very similar to the first. The inaccurate profiles have lost some short wavelength content. This suggests that the oil accumulates locally and that it is thick enough at the interface not to settle

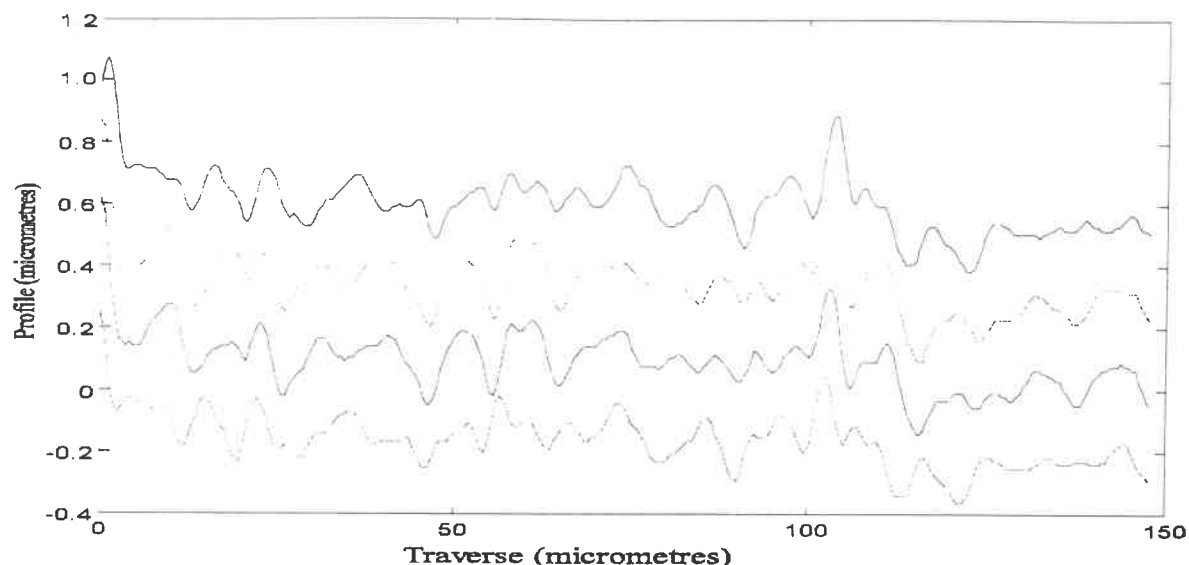


Figure 3: Successive profiles (from top) on a flooded surface, lowest after lifting off surface.

naturally in a period of seconds, but a reverse traverse will clear it. Perhaps the oil distribution on the tip has an effect on fidelity since there is time between traces for oil to flow back onto surface that is relatively distant from the contact point.

Conclusions. Conventional surface profilometry appears to have low sensitivity to lubrication conditions at the tip. We have found no convincing evidence of changes in profile fidelity between clean surfaces and those flooded or wiped with mineral oils of viscosities typically used for light lubrication, except at the initial application of a flood. Friction coefficients at the tip were largely independent of the amount of lubricant present, perhaps because there is at best boundary lubrication at the actual tip under the very high local pressures that occur there. Lubrication may offer a little protection against surface damage of softer materials at normal stylus forces.

Lubrication does not give a low-cost means of usefully improving instrument performance, and flooding with oil may reduce throughput because of the need to settle the trace. On the positive side, even as more stringent demands are placed on surface metrology systems, there will be little need for close control of cleaning on components typical of the metalworking industries. Also, accurate measurements of surfaces that are best protected from the atmosphere may be possible through a complete covering of mineral oil. If this is done an initial trace must be taken and discarded in order to settle the system into its standard operational conditions.

References.

- [1] X Liu, D G Chetwynd, S T Smith & W Wang: Improvement of the fidelity of surface measurement by active damping control. *Meas. Sci. Technol.* **4**, 1993, 1330-1340
- [2] X Liu, S T Smith & D G Chetwynd: Frictional forces between a diamond stylus and specimens at low load. *Wear* **157**, 1992, 279-294
- [3] B J Hamrock: *Fundamentals of Fluid Film Lubrication*. 1994 New York: McGraw Hill

Acknowledgment

We are grateful to EPSRC for their support of this work as part of grant GR/K 71332.

Development of Dual Servo STM

Kimiyuki MITSUI* and Takayuki GOTO**

* Department of Mechanical Engineering, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223 Japan

** Advanced Technology Research Center, Mitsubishi Heavy Industries, Ltd.
1-8-1, Sachiura, Kanazawa-ku, Yokohama, 236, Japan

Introduction

Precise surface roughness measuring techniques are becoming increasingly important in many industrial fields such as ultra-precision machining. Moreover new measuring techniques for forms and dimensions accuracy of micro parts which compose micro machines are strongly demanded.

For the measurements of precision surfaces, instruments with diamond stylus, optical surface roughness measuring methods^{1,2} and STM (Scanning Tunneling Microscope)³ have been developed. Optical measuring methods have advantages of non-contact measurement, that is, no fear of making scratch on the measured surface and they normally have high measuring speed. But their measuring range in the direction vertical are not wide. STM is also very effective method for surface roughness measurement, owing to its ultra-high lateral and vertical resolution. But STM is not sufficiently useful for surface roughness measurement of machined parts, because of the small measuring range especially in the direction vertical.

From these points of view, a new surface profile measuring instrument that has high resolution and wide measuring range has been developed. Newly developed dual servo STM has a PZT driven double-compound parallel springs with 60 μ m travel for coarse mechanism and a PZT with 15 μ m travel for fine mechanism. The coarse and fine mechanism are connected in series, and a thin probe is attached at the fine mechanism PZT. Both coarse and fine mechanism are controlled simultaneously to get stable tunneling current between measured surface and probe.

Dual servo control technique have been applied for high precision mechanical stages which require high positioning accuracy and long travel^{4,5}, but the application of the dual servo control technique to the surface profile measurement has not been reported.

In this report measuring principle and basic characteristics of the dual servo STM are described. Then it is used to measure surface roughness of machined workpieces and the results are compared with those obtained by a conventional stylus method.

Measuring principles and system layout of the apparatus

Figure1 illustrates configuration of the dual servo STM. The apparatus comprises a

sensing probe that detects tunnel current, coarse and fine mechanisms that actuate the probe and a electrical circuits for controlling moving mechanisms. The coarse mechanism comprises two parallel springs with flexure pivoted arms in series⁶, and displacement of two PZT (maximum displacement $15\mu\text{m}$) is amplified up to $60\mu\text{m}$. The fine mechanism is a PZT of maximum displacement $15\mu\text{m}$ which attached at the end of the coarse mechanism. A tungsten probe attached to fine mechanism was used as the tunnel current detecting probe. Surface profile is measured by controlling gap between the probe and target surface so that the stable tunnel current is obtained, while target surface is moved by X axis feed gear.

Displacement of the coarse mechanism is measured by a capacitance type displacement sensor. On the other hand, direct displacement measurement for fine mechanism is difficult. Therefore observational PZT for fine mechanism is used to estimate the displacement of it. Summation of both fine and coarse mechanism displacement gives displacement of the probe.

Figure2 shows a block diagram of dual servo control system. Here, K is a proportional constant relating to I-V converter which detects tunnel current and logarithmic converter. R_f and R_c represent reference input for fine and coarse mechanism respectively. D_f and D_c express displacement of fine and coarse mechanism. Tunnel current detected by probe is converted to voltage signal proportional to the gap between surface and probe. The

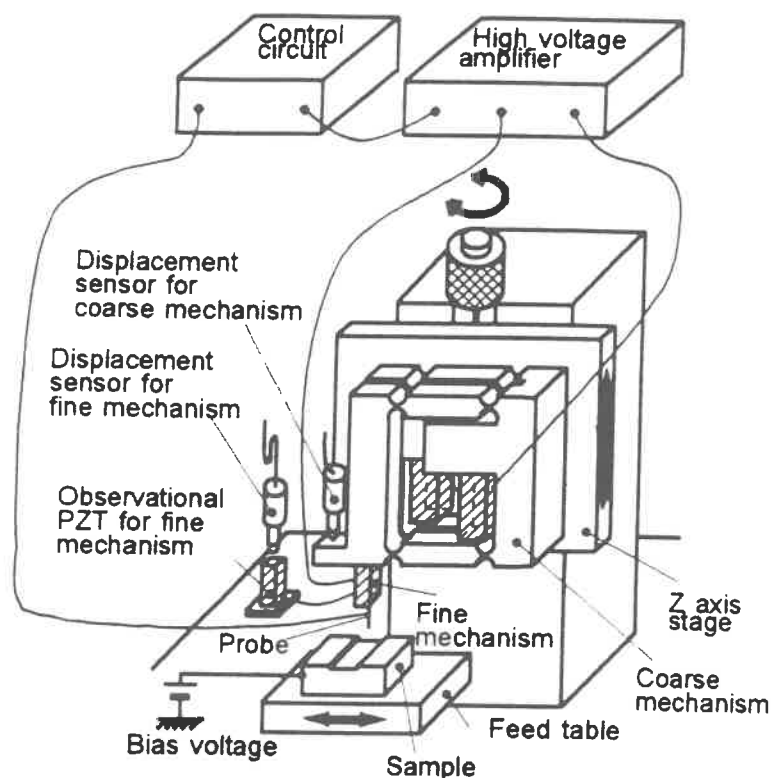


Fig. 1 System layout of the apparatus

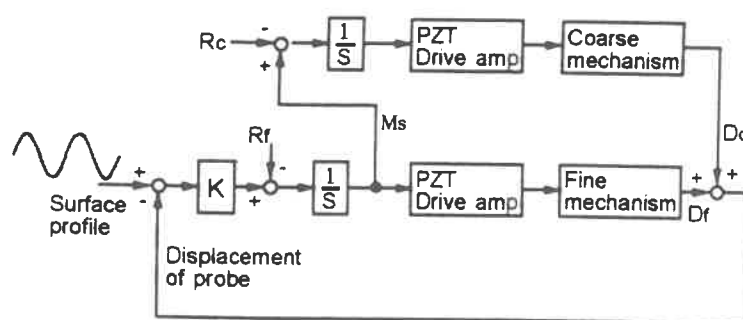


Fig. 2 Block diagram of dual servo control system

difference of this voltage and reference input for the fine mechanism R_f is integrated by integrator, and the signal is fed to fine mechanism through PZT drive high voltage amplifier. According to the displacement of PZT for fine mechanism is controlled so that the voltage difference of R_f and logarithmic converter output becomes zero, the distance between target surface and probe is always kept constant. On the other hand, integrator output M_s is given also to control circuit for coarse mechanism. Voltage difference of M_s and reference input R_c for coarse mechanism is integrated by integrator and displacement of coarse mechanism is controlled. The PZT for fine mechanism stretches $15\mu\text{m}$ when 100V is applied to it, so the value R_c is established as 5V in consideration of magnification 10 of PZT drive amplifier. As a result of this, the displacement of coarse mechanism is controlled so as to the displacement of fine mechanism always keeps $7.5\mu\text{m}$.

Measuring results

To illustrate the broad range of measurements that can be performed with this instrument, several specimens with different surface features and dimension ranges were measured, and the results were compared with those obtained by a stylus method. All the profiles here are sampled with measuring speed of 0.03mm/s and recorded by FFT analyzer without using both analog and digital low-pass filter.

The first measurement was performed on a steel block with rectangular groove $10\mu\text{m}$ deep. Figure 3 compares the dual servo STM measurement with a diamond stylus. The agreement of these two instruments is excellent, but in the result obtained with the dual servo STM method, higher frequency components are observed.

Figure 4 shows surface roughness profile of a milled aluminum plate. Cutting was performed with spindle speed 91rpm and feed rate of 80mm/min . It can be seen that the two measuring results are in reasonably close agreement.

Figure 5 shows also surface roughness profile of a milled aluminum plate. Cutting was performed with spindle speed 67rpm and feed rate of 80mm/min . The approximate peak-to-valley height of the profile $30\mu\text{m}$ exceeds maximum displacement of fine mechanism, so the validity of dual servo STM which works as the combination of both coarse and fine mechanism, was shown. The width of peak obtained by stylus method is wider than that obtained by dual servo STM. The reason for this depends on probe radius of both stylus and STM method.

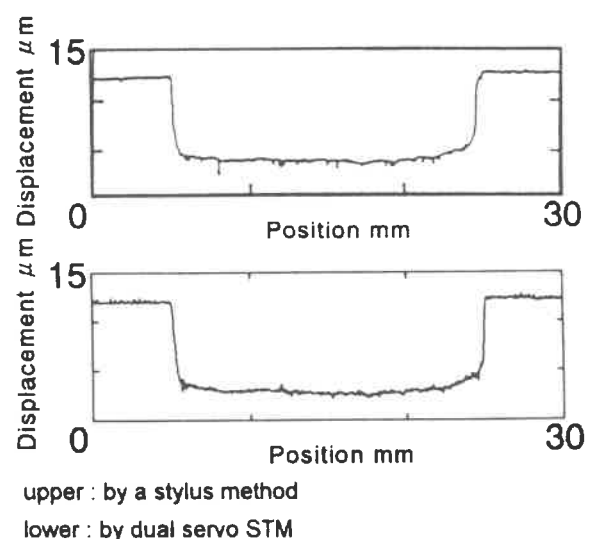
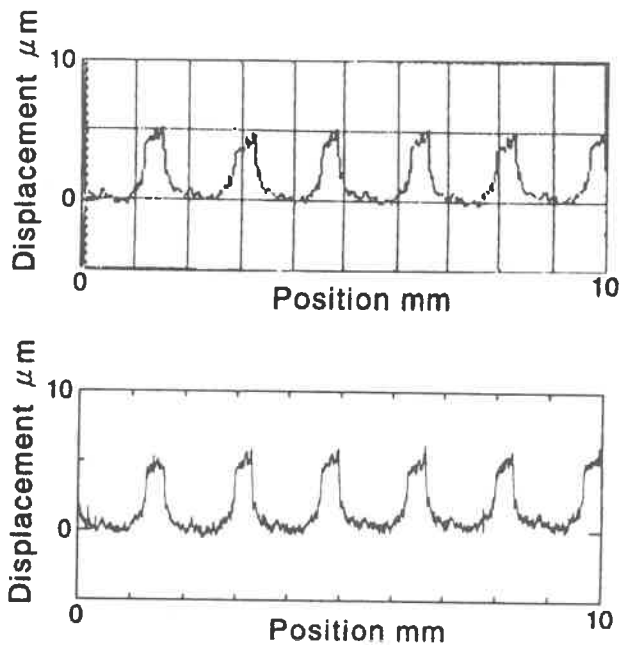
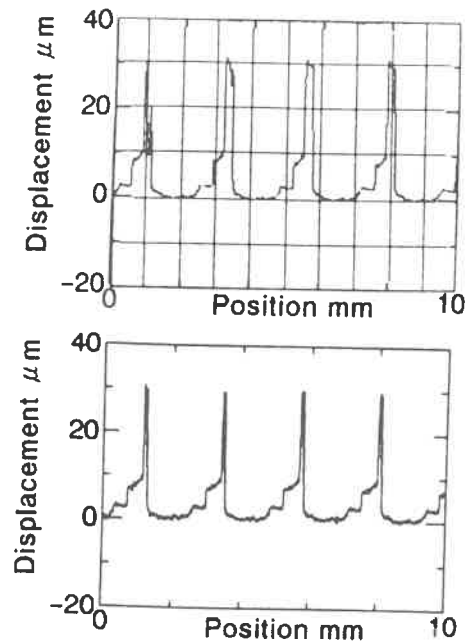


Fig. 3 Surface profile of the groove



upper : by a stylus method
lower : by dual servo STM

Fig. 4 Surface profile of a milled workpiece (1)



upper : by a stylus method
lower : by dual servo STM

Fig. 5 Surface profile of a milled workpiece (2)

Conclusions

A newly developed dual servo STM with coarse and fine mechanism and consistent control algorithm has been described and evaluated. The results are summarized as follows.

- 1) By the dual servo STM, stable tunneling current between measured surface and probe can be obtained.
- 2) Surface profile of several specimens were measured by the new instrument, and the results obtained agree well with that obtained by a stylus method.

References

- 1) T. Kohno et al: High precision optical surface sensor, *Applied Optics*, 27-1(1988)103.
- 2) K. Mitsui, M. Sakai and Y. Kizuka: Development of a high resolution sensor for surface roughness, *Optical Engineering*, 27-6(1988)498.
- 3) G. Binnig et al: Surface Studies by Scanning Tunneling Microscopy, *Phys. Rev. Lett.*, 49-1(1982)57.
- 4) S. Sakuta, K. Ogawa and K. Ueda: Experimental Studies on Ultra-precision Positioning - Inchworm Movement Method using Fine and Coarse Positioning -, *J. of the JSPE*, 57-12(1991)2133. (in Japanese)
- 5) Y. Okazaki, S. Asano and T. Goto: Dual-Servo Mechanical Stage for Continuous Positioning, *Int. J. of the JSPE*, 27-2(1993)172.
- 6) N. Tsuda et al: Wide Range STM, *J. of the JSPE*, 55-1(1989)146. (in Japanese)

Direct Comparative Length Measurement between Regular Crystalline Lattice and SEM Standard Grating Using Dual Tunneling Unit STM

*Masato AKETAGAWA, *Koji TAKADA, *Shigetoshi SASAKI, *Seiki SUZUKI,
*Keiko KOBAYASHI, *Kazutoshi YAMADA, **Yoshinori NAKAYAMA

*Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata 940-21, JAPAN
**Central Research Laboratory, Hitachi Ltd., 1-280 Hiagshi-Koigakubo, Kokubunji, Tokyo 187, JAPAN

I. Introduction

As research activities in the fields of nanotechnology expand, a new method is required to measure lengths of a few micrometers precisely with sub-nanometer resolution. At present, the industrial standard of precise length measurement is the wavelength of a laser, and laser interferometry is widely used. However, it is very difficult to obtain sufficient accuracy in length measurements with sub-nanometer resolution using a laser interferometer in air.¹ Therefore, in the nanotechnology field, a new length measurement method applicable to the range of a few micrometers with sub-nanometer resolution is required.

The lattice spacings of some regular crystallines, which are about 0.2 nm, are stable and uniform over a long range, sometimes a few micrometers, when the crystallines are stress-free. Therefore, these crystallines can be used as reference scales with sub-nanometer resolution. On the other hand, a scanning tunneling microscope (STM)² can be used to obtain images of the atoms in a crystalline surface in air. It is, therefore, expected that a "digital lattice scale" with sub-nanometer resolution can be obtained using a crystalline as a reference scale and the STM as a detector.^{3,4} We have constructed an instrument for length measurement using a dual tunneling unit STM (DTU-STM).^{4,5} The DTU-STM consists of one X-Y scanning table and two tunneling units A and B controlled in the Z-axis direction independently. The sample to be measured and a reference crystalline are set on the X-Y table immediately beneath the tunneling units A and B, respectively. If the relative motion of the X-Y table with respect to the tip of the tunneling unit A is the same of the tunneling unit B, simultaneous STM imaging with both units enables precise comparative length measurement.

In this paper, we consider 1) feasibility evaluation of the comparative length measurement over 1 micrometer using a crystalline lattice (highly oriented pyrolytic graphite-HOPG)

and the DTU-STM, and 2) direct comparison between the SEM standard grating pitch, which is 240 nanometers, and HOPG lattice spacing using the DTU-STM.

II. Apparatus and method for correcting measurement error

Figure 1 shows schematically the configuration of the DTU-STM used in the experiments. A tube-type piezoelectric actuator is used as the X-Y scanner. To reduce thermal deformation effects, the STM body was designed to be almost cylindrically symmetric around the vertical center line of the X-Y scanner and super-invar (low-linear-expansion steel) is used for the body. Z-axis control of the tunneling units A and B is carried out using a modified NanoScope III and an auxiliary STM controller, respectively.

The dimensional errors of the present comparative length measurement method are mainly due to a) the difference of the tilting of the measured surface plane of the two tunneling units (tilting error), and b) the difference of thermal drift of the two tunneling units (thermal drift error). The tilting error can easily be corrected using the tilting angle between the X-Y scanning plane and the average plane of the measured surface. The tilting angle can be de-

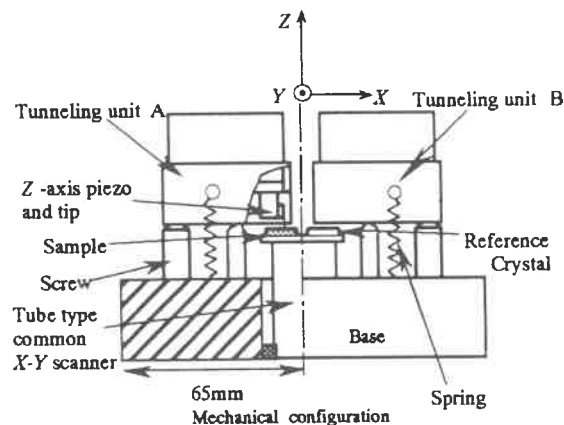


Fig. 1 Experimental DTU-STM (side view)

terminated directly by applying the least squares method to the three-dimensional STM image data. In the following discussion, we assume the tilting error has been calibrated using the above method.

The thermal drift error for lattice vector, $\delta \mathbf{a}_i$, is expressed by the equation⁵

$$\delta \mathbf{a}_i = \mathbf{a}_i' - \mathbf{a}_i = t(\mathbf{a}_i') \mathbf{d} \quad (i=1,2) \quad (1)$$

where $\mathbf{a}_i'$ and $\mathbf{a}_i$ are a lattice vector with and without thermal drift, $\mathbf{d}$ is thermal drift velocity vector, and $t(\mathbf{a}_i')$ is the period required to scan from the origin to the end point of $\mathbf{a}_i'$ vector. If the thermal drift doesn't change during measurements, $\mathbf{d}$ can be determined by comparing a pair of corresponding lattice vectors in two STM images obtained at different scanning rate. $\mathbf{d}$ is expressed as⁵

$$\mathbf{d} = \frac{\mathbf{a}_{if}' - \mathbf{a}_{is}'}{t_{if}(\mathbf{a}_{if}') - t_{is}(\mathbf{a}_{is}')} \quad (i=1,2) \quad (2)$$

where $\mathbf{a}_{if}'$ or $\mathbf{a}_{is}'$ is lattice vector in fast- and slow-scanned STM images, and $t_{if}(\mathbf{a}_{if}')$ or $t_{is}(\mathbf{a}_{is}')$ is the scanning time for $\mathbf{a}_{if}'$ or $\mathbf{a}_{is}'$. As averaged lattice vectors are determined by 2 dimensional FFT analysis,⁶ therefore the comparison of a pair of lattice vectors can be made in the spatial frequency domain.⁵

III. Experimental results and discussion

(A) Feasibility Test

To verify the feasibility of the proposed method for micrometer range length measurement, two samples of highly oriented pyrolytic graphite (HOPG) were set on the X-Y table immediately beneath the tunneling units A and B, and the measurement accuracy was examined. Length measurement was performed along the Y- (slow scanning) axis, and displacement of Y-scanner determined from the image of the two units was compared. The displacement of the Y-scanner can be determined by counting the number of lattice points in the STM image along the Y-axis, but this displacement includes the thermal drift error. To detect this error, a modified ramp signal for Y-axis scanning, which is shown in Fig. 2, was utilized. The modified ramp signal provides fast and slow scanning parts alternately and each part forms a unit STM frame of 256×256 pixels. Comparing the image of a slow or fast scanning part with that of the neighboring fast or slow part, the local drift velocity vector $\mathbf{d}$ can be obtained using the above method. Then, the displacement of the Y-scanner can be determined accurately using drift velocity vector $\mathbf{d}$.

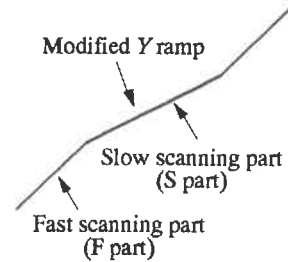


Fig. 2 Modified Y ramp signal

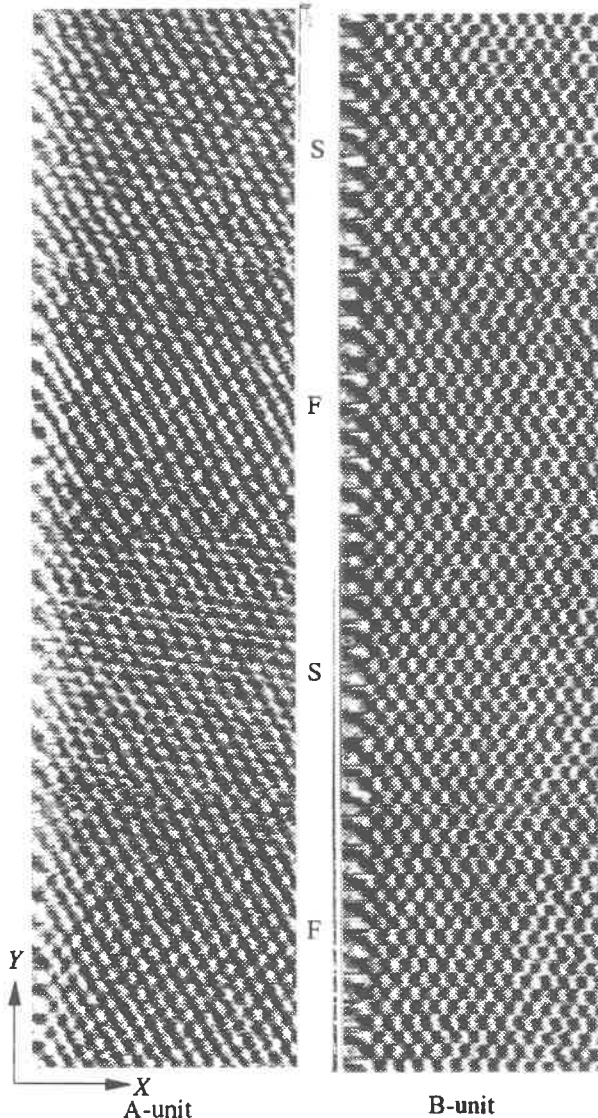


Fig. 3 Several frames of complete image of HOPG
Scanned range (4 nm (X) × 1 μm (Y))
F: fast scanning part, S: slow scanning part

Figure 3 shows several unit frames in the raw STM images of a region of HOPG about 1 μm long obtained simultaneously using the tunneling units A and B. The complete image of

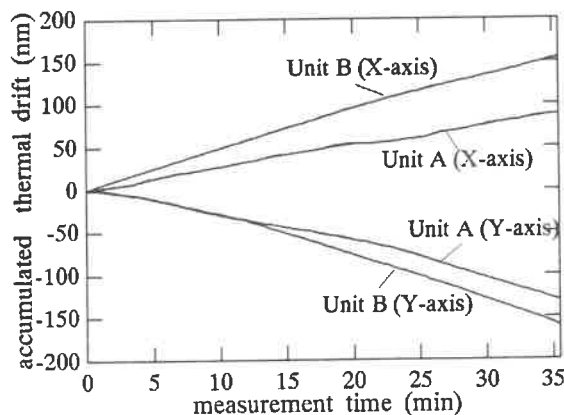


Fig. 4 Relationship between measurement time and accumulated thermal drift

the 1- μm -long region consists of a continuous series of 256 unit frames. It took about 36 min. to obtain the complete image. In the complete image of 1- μm -long region, no atomic step and no lattice defect are detected, which verifies the feasibility of using HOPG as a practical reference scale in the proposed comparative length measurement.

Figure 4 shows the relationship between the measurement time and the accumulated thermal drifts of the two units, which are obtained from thermal drift velocities using the method described in Section II. The relationship between the voltage applied to the scanner and the displacements of the tunneling units in the Y direction is shown in Fig. 5. After correcting the errors due to the thermal drifts given in Fig. 4, the measured displacements of Y scanner from the images of the units A and B were found to coincide with each other. The maximum difference between the two was less than 8 nm over the measurement length of 1 μm .

(B) Direct comparative length measurement between crystalline lattice and SEM standard grating

SEM standard grating and HOPG were set on X-Y table just beneath tunneling units A and B, and images of both the grating and the HOPG were simultaneously obtained in the range over 1 micrometer using the DTU-STM. A pitch of the grating, which is 240 nm, was calibrated by the diffraction method.⁷ Figure 6 shows the SEM image of the grating. The grating coated with Au was used in the experiment. Length measurement was also performed along the Y- (slow scanning) axis. To reduce thermal drift error, a rapid scanning method, with which fast (X-) axis scanning frequency is 500 Hz, was utilized for the experiment. The scanned time for the complete image was 130 seconds. Figure

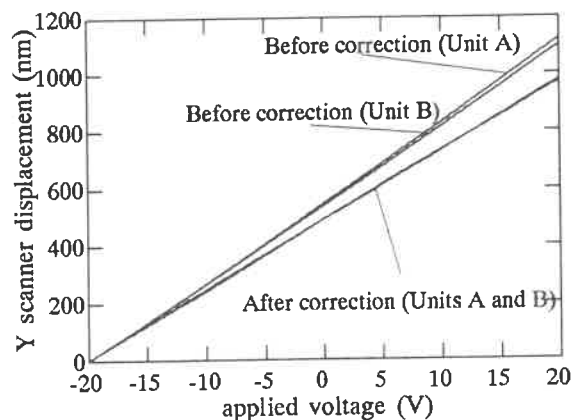


Fig. 5 Relationship between applied voltage and displacement of Y scanner

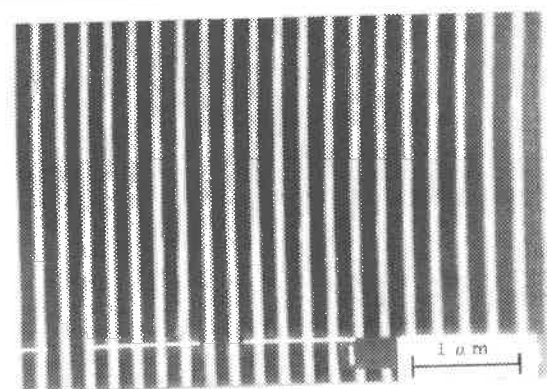


Fig. 6 SEM photograph of the SEM standard grating (pitch = 240 nm)

7 shows several frames of HOPG atomic image and a cross section of the grating. In the complete HOPG atomic image, there were also no atomic defect and step. The thermal drift error of both the HOPG and the grating images was less than 14 nm for the measurement length of 1 micrometer. The averaged value and the standard deviation of the pitch are 237 nm and 23 nm, respectively. We observed just 5 pitches of grating, but measured value of the pitch was very closed to the calibrated value within a error of 2 %. We believed this error was introduced by the remaining thermal drift error and corrugations of the grating.

IV. Conclusion

(A) In comparative length measurements using DTU-STM, the relative motion of the X-Y table with respect to the tip of the tunneling unit A should be the same as that of the tunneling unit B. In the first test, two sample crystallines of highly oriented pyrolytic graphite (HOPG) were set on the X-Y table immediately beneath the tunneling units A and B, and the displacement

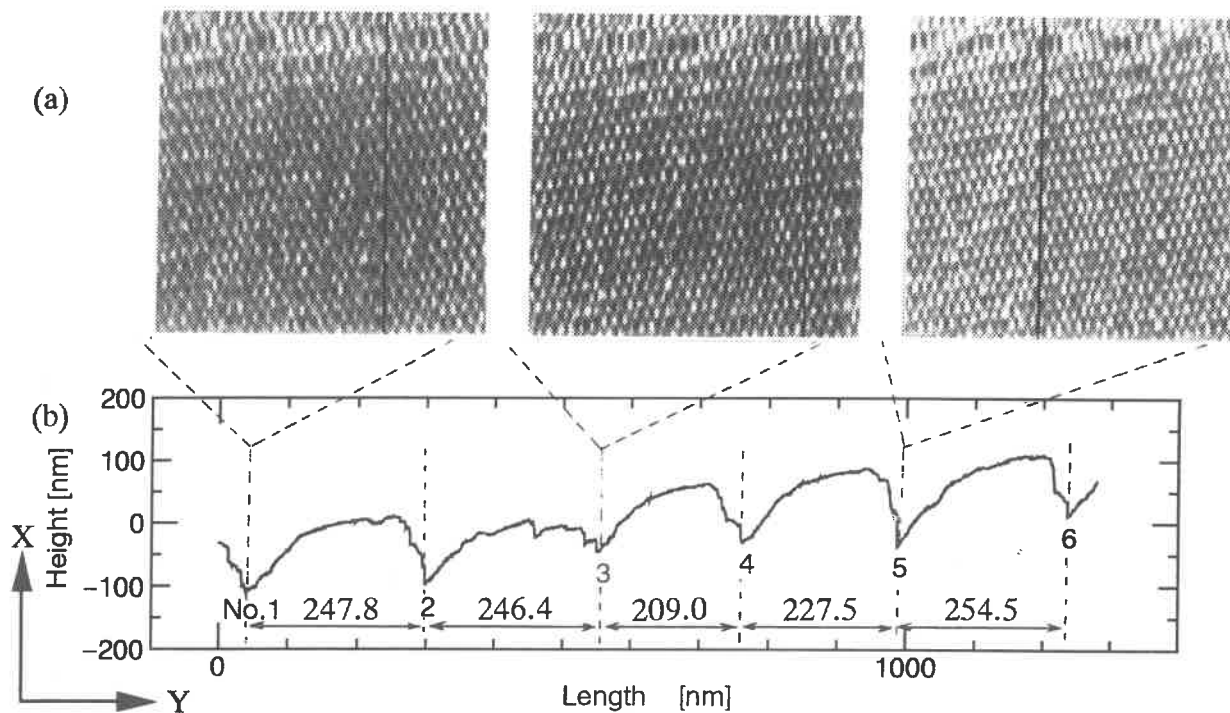


Fig. 7 (a) HOPG atomic image and (b) Cross section of the grating

of Y table measured from the images of the two units were compared in the range of 1 micrometer. In the test, very fine atomic images of about 1- μm - long region, without any lattice defects or irregular steps, were successfully obtained from both of the units. This result shows that HOPG can be used as the reference crystalline. After correcting thermal drift error by the present method, both of the measured displacements of Y table from the images of units A and B were coincided with each other within an error of 1 % over the measured length of 1 μm .

(B) In the second test, the SEM standard grating, whose pitch is 240 nm, and HOPG were set on X-Y table just beneath tunneling units A and B, and images of both the grating and the HOPG were simultaneously obtained in the range over 1 micrometer using the DTU-STM. To reduce thermal drift error, a rapid scanning, with which fast axis scanning frequency is 500 Hz, was utilized for the experiment. The

measured value of the pitch using present method was very closed to the value calibrated by the diffraction method within a error of 2 %.

References

- ¹ R. Thalmann and W. Hou, Proceedings of the 7th IPES, Kobe, Japan, 11(1993)
- ² G. Binning, H. Rohrer, CH. Gerber and E. Weibel, Appl. Phys. Lett. **40**, 178(1982)
- ³ E. C. Teague, J. Vac. Sci. Technol. **B7**, 1898(1989)
- ⁴ H. Kawakatsu, Y. Hoshi, T. Higuchi, and H. Kitano, J. Vac. Sci. Technol. **B9**, 651(1991)
- ⁵ M. Aketagawa and K. Takada, Nanotechnology **6**, 105 (1995)
- ⁶ C. Kittel, "Introduction to Solid State Physics", John Wiley & Sons(1976)
- ⁷ Y. Nakayama, S. Okazaki and A. Sugimoto, J. Vac. Sci. Technol. **B6**, 1930(1988)

CYLINDRICITY - MORE THAN JUST A NUMBER ?

Philip J. Castle
Rank Taylor Hobson Inc
Rolling Meadows, IL

Introduction

A recent addition to the science of roundness measurement is cylindricity. With this parameter, it is possible to generate one number to completely describe a cylindrical component. However, cylinders with various differing characteristics can have the same cylindricity value, despite the fact that these characteristics can have a marked effect on the function and performance of the part. This presentation will teach you how to measure cylindricity (and some of the pitfalls to avoid), describe the three features that comprise a cylinder, and the effect on function of these features.

Roundness and cylindricity are measurements of the variation in geometrical form of nominally cylindrical objects. In order to totally assess such an item dimensionally, it must be measured for several distinct characteristics - geometrical form (roundness, cylindricity), surface finish (parameters such as Ra), size/dimensions and locations. Surface finish is the province of a surface finish gage, absolute size, locations and dimensions are typically the province of devices such as CMMs.

A simple example of this would be a soft drink can. This must have a certain surface finish for successful painting, a certain round or cylindrical form in order for the end to fit properly, and be of certain dimensions to fit into filling and vending machines. While current inspection instruments, particularly higher priced ones, have some overlap in these functions, there are generally compromises or trade-offs involved. Also, there can be interactions - an out of roundness condition may 'use up' part of the size tolerance, and surface finish may affect the roundness value - although filters are normally chosen to minimize this particular effect.

Roundness Measurement Techniques

Cylindricity results are derived from a set of roundness measurements. There are three principle methods of assessment of out-of-roundness. Each has its advantages and disadvantages.

Diametral, such as using a micrometer. Here, the part is measured for diameter at least two locations. The disadvantage of this approach is that it will not detect odd numbered lobes. For example, a three lobed part could appear to be constant diameter, and therefore

erroneously considered to be 'round'. One of the fundamental weaknesses of this method is that it does not employ an axis of rotation.

Chordal, such as indicator and V-block or centers. Here, a dial indicator is fixed in position, and the part is rotated, in a V-block or between centers, under its point. While this does introduce a rotation, it is referenced to the outside diameter of the part. Again certain lobing will be masked or exaggerated, depending on the relationship of their wavelength to the included angle of the v-block.

Radial, such as with a roundness gage. The principal here is that the part to be measured rotates around an axis, ideally its own center, and a probe measures variations in radius around the part. This is the basis of most current roundness gages, and is the configuration around which current standards are written.

Standards

Standards are the 'rules' by which businesses ensure that they are following common procedures. They also, to an extent, define optimum, good quality, methods. The US standard for roundness measurement is ANSI B89.3.1, 'Measurement of Out-Of-Roundness'. Amongst other topics, this Standard defines two of the cornerstones of good roundness and therefore cylindricity measurement, these being reference figures and filters.

The above two methods will yield numbers which may bear some relationship with the actual roundness of the part, and a manufacturer may be able to develop a correlation between such results and the performance of the part. However, neither will yield results which could be described with confidence as the true out-of-roundness of the part. The existence of the two 'rules' given above means that the chordal and diametral methods described can seldom, if ever, yield results which comply with ANSI standards.

Reference figures.

In order to assess roundness, a reference circle (figure) must be fit to the measurement. This defines the position of the center of the measurement, which in turn is used to define the out of roundness value. Also, these centers may be used to define an axis, for concentricity, coaxiality and squareness assessment. There are four reference figures, each of which has a particular application. The two most commonly used follow.

1. Minimum radial separation. This is two concentric circles drawn around the data (not necessarily the sum of the MI and MC circles, as these are unlikely to be concentric), one inside and one outside, with minimum separation which totally enclose the measurement

data. The roundness value is the difference in radius of these two circles, hence the term Minimum Radial Separation. This figure is also called Minimum Zone. Of the four figures, this will always give the lowest ('best') roundness result. This is the figure which the roundness standards default to, and is analogous with the standard GD&T definition of circularity.

2. Least squares. This is a mathematical, rather than functional, fit, being a circle which is positioned such that the area of the peaks outside the circle are equal to the area of the valleys inside the circle. This is by far the most stable center, and is often used in preference to the MRS figure for this reason.

Filters

Filters are used to remove the surface finish of a part, mathematically, to reveal the underlying round and cylindrical form. These are normally expressed in terms of UPR, Undulations Per Revolution, and remove higher UPR effects (high frequency surface finish) of the measurement.

Cylindricity

Cylindricity is the nett effect of several (at least two, typically as many as twenty) roundness planes measured around the cylindrical component, normally equi-spaced vertically. The vertical motion of the probe between successive measurements must be both straight and parallel to the axis of rotation of the measuring instrument. Any error in these two parameters will add to the uncertainty in the cylindricity result.

What do we mean by 'nett effect' ? Suppose we have a roundness instrument with a vertical motion which is straight, typically to 10 or 20 uin, (achieved either purely through high mechanical precision, or high precision combined with software correction) and perfectly aligned with the spindle axis of rotation. If we take several roundness measurements at different heights with such an instrument, it can be said with confidence that any variations in the positions of the centers, or in the size of the planes, are truly in the part itself (and not due to variations in the vertical motion). This being the case, it is possible to plot each of the planes in their measured relative positions and sizes. Then, we can draw two coaxial cylinders with minimum separation around the measured planes, one around the inside and one around the outside, which just enclose all of the measurement data. The difference in radius (the Minimum Radial Separation) of these two cylinders is the cylindricity value. This cylindricity result, then, is the nett effect of the roundness in each individual plane, *plus* the error in the alignment of the centers ('banana' shape) *plus* the variation in size (taper, 'barrel' or 'apple core').

Part Measurement

Examples of parts where roundness alone may not be enough, and where a cylindricity inspection would be required, are calibration artifacts such as gage pins and ring gages, and manufactured parts include roller bearings, fuel injector and hydraulic valves, and piston wrist pins. A cylinder can be an OD surface, such as a gage pin, or an ID surface, such as a bore.

Consider a sample of three parts which would normally be expected to have good cylindricity, such as a hydraulic valve component. We want to check each of these for cylindricity. Lets also say that we have a modern, high precision cylindricity gage to measure the parts on.

Our three cylindricity results turn out to be the same, perhaps 50 millionths or so. However, when we look at the three graphical results displays, they appear to be rather different. The first is essentially straight, with no apparent taper or other change in size. However, the surface of the part is relatively rough. The second part does not exhibit any of the rough surface seen the first one, and it does not appear to change in size - ie, no taper. What it does have, though, is a bowed ('banana') shape. The third part has a similarly good surface, is not bowed, but does appear to be tapered. Due to the way cylindricity results are computed, as described above, it is certainly possible for a group of parts such as these to have rather different forms yet still have the same cylindricity value.

What will tend to happen is that the cylindricity tolerance applied to the part will be set sufficiently tight to cover any of, or a combination of, or all of these effects.

Consequently, the tolerance may be much tighter - making the part unnecessarily expensive to manufacture than is really necessary for successful function of the part. For example, in the hydraulic example, the main concern may be leakage, which would be largely governed by the surface of the part. Bowing would be of a lesser concern, and taper would be of the least concern. However, we may be applying a tight cylindricity tolerance to ensure that the surface is within limits, and consequently are expending unnecessary effort making parts with much lower taper than is really acceptable.

There are many components where cylindricity type measurement is essential, for example wrist pins, hydraulic component, fuel systems, prosthetics and bearings. Consideration of the part's function should include whether the part is for a rotating or static application. Furthermore, the surface finish and size of the part are often also important considerations, and will have an impact on the performance of the part also.

OPTICAL POLYGON CALIBRATION WITH A He-Ne LASER

Rufino Díaz, Rigoberto Nava

Centro de Instrumentos, UNAM, Apdo. Postal 70-186, Coyoacán . 04510, México, D.F.
Tel. (525) 550-0416. FAX (525) 550-0357. E-mail: rufino@aleph.cinstrum.unam.mx,
navar@aleph.cinstrum.unam.mx

ABSTRACT

A method for measuring the angular variations of the faces of an optical polygon without using autocollimators is proposed. The method only uses one He-Ne laser as a light source; the beam is divided into two beams in order to sense by autocollimation two adjacent faces of the polygon, each at a time. As usual, a positive lens is used to translate the angular variations of the reflected beams to transversal displacements on the focal plane. The experimental set up and procedure and the theoretical basis of the method are described. It is experimentally shown that the method allows one to test the polygon with an error less than 3 arc sec.; however this is not a fundamental limit.

1. INTRODUCTION

Optical polygons are employed as angular patterns which divide circumference in N equal parts (N is the number of faces). This allows one, for example, to calibrate angular divisor turntables used for gear manufacturing or for measuring gear pitch. In order to measure angles with high precision, the optical polygon geometry should have narrow tolerances on the angle of each of its faces.

As every reference pattern, the optical polygon must be calibrated. That is, the error in the angular orientation of each of its faces, referred to one of them, must be measured; the reference face is called the zero degrees face (0°).

Usually, the polygon is calibrated with the help of two autocollimators [2]. Roughly, the method consists in aligning the optical axis of the autocollimators with the normals of two adjacent faces. Then, the angle between each other pair of adjacent faces with the angle of the autocollimators is measured. Due to the fact that the sum of all the angles between adjacent faces must be 360° , no matter how big the errors are, then the sum of all errors must vanish. This allows one to compute the error for each face.

Often, however, there are not available two autocollimators; some cases one does not have even only one at hand. For that reason, in this paper, a method for measuring the angular variations of the optical polygon faces without using autocollimators is proposed. The method makes use only of one He-Ne laser as a light source; the beam is divided into two beams in order to sense by autocollimation two adjacent faces of the polygon, each face at a time. As usual, a positive lens is used to translate the angular variations of the reflected beams to transversal displacements on the focal plane.

In the next sections, the theoretical basis of the method is presented, as well as the experimental set up and some experimental results.

2. METHOD

2.1. Experimental set up

The two autocollimators are substituted by the laser beam as the directional reference, but in order to use only one laser, the original beam is divided into two beams with the help of a beam splitter (Fig. 1); one beam, the transmitted one, is directed to one of the faces of the polygon whereas the reflected one, is reflected again on an auxiliary optical flat mirror to be sent to an adjacent face; in the following text, those beams will be called beam 1 and 2, respectively. Each beam is normally incident on the corresponding face, so that after reflection each beam goes back along the same path. Now, at the beam splitter, we consider the reflected beam of the beam 1 and the transmitted one of the beam 2; from this point they follow almost exactly the same path. A positive lens of large focal distance is placed in the path of the beams and, in its focal plane, a knife edge blocks out around a half of each beam. Just behind the last, a photodetector measures the transmitted power for each beam; this gives an accurate measure of the transverse beam displacements [3]; these are related to the deviations of the beams produced by the angular errors in the faces of the polygon. In order to measure one beam at a time, the screens $p1$ and $p2$ are used to blocking the beams in an alternative way.

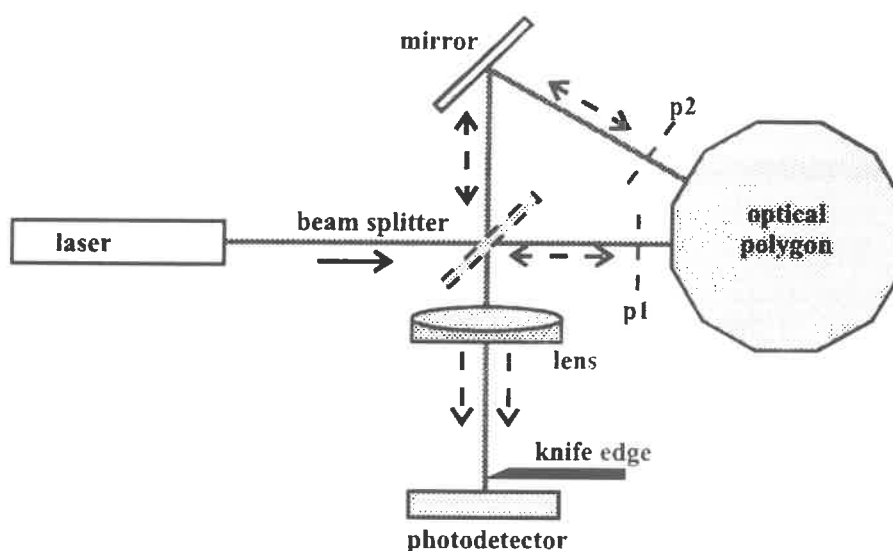


Figure 1. Experimental set up to calibrate optical polygons by using only one laser.

2.2 Experimental procedure

Once the set up is aligned when the beams point towards the face 0 and 30 degrees, the radius of each beam is measured (without the knife edge). Then, blocking one beam with a screen, for instance the beam 2, the total power of the beam 1 is measured. This can be done by rotating the polygon a small angle to avoid the beam being blocked by the knife and arriving at the detector with the full power. Now, the polygon is returned to the position where the beam is half blocked by the knife; the transmitted power is measured again. Next, the beam 1 is blocked by the screen $p1$ and the transmitted power for the beam 2 is measured, without changing the orientation of the polygon; the total power of the beam is also measured. This procedure is repeated for the

other pairs of adjacent faces, in each case the polygon must be oriented in such a way that the transmitted power of the beam 1 is almost half its total power.

2.3 Formulas for computing the angular errors.

2.3.1 Beam radius.

In the literature there are reported several methods for measuring the radius of a gaussian beam (see ref. [3] and refs. there in). Probably one of the simplest is the opaque wire method [4], because one only needs to blocking out the beam with a calibrated opaque wire of thickness Δ and bring it to a transversal position that yields the minimum transmitted power P_o . If P_T is the total power of the beam, its radius r_o at the knife edge position is given by

$$r_o = \frac{\Delta}{2} \frac{1}{\text{erf}^{-1}\left(1 - \frac{P_o}{P_T}\right)}, \quad (1)$$

where erf^{-1} is the inverse of the well known error function.

2.3.2 Transversal displacement of the beam

Accordingly with R. Díaz-Urbe, *et al* [3], the relative position x , between the knife edge and the center of a gaussian beam is given by

$$x = \frac{r_o}{\sqrt{2}} \text{erf}^{-1}\left[\frac{2P}{P_T} - 1\right], \quad (2)$$

where P is the transmitted power of the beam when is partially blocked by the knife edge. Eq. (2) does not involves approximations, and is useful for $|x| \leq r_o$. Furthermore, for very small values of x , it is possible to make the next approximation

$$x = \sqrt{\frac{\pi}{8}} r_o \left[\frac{2P}{P_T} - 1\right]. \quad (3)$$

Such an approximation does not differs more than 1% of the true value if $|P - P_T| \leq P_T/20$.

2.3.3 Angular Displacement

From Geometrical Optics is well known that a ray arriving at a positive lens making an angle α from the optical axis of the lens, emerges in such a way that in the focal plane attains a height x from the same axis, given by

$$\tan \alpha = \frac{x}{f}, \quad (4)$$

where f is the focal distance of the lens; for small angles one can make the next approximation: $\tan \alpha \approx \alpha$. Eq. (4) allows one to compute the angle of the beams 1 and 2, but it must be kept in mind that a reflected ray from a mirror rotates an angle 2θ , when the mirror is rotated an angle θ , so the angle variation of the faces of the polygon is twice the angle of each corresponding beam.

2.3.4 Angular variation of the polygon faces

The theory to calibrate the polygon is described by Hume [1] and Nava [2], the formulation proposed by the last will be followed here.

Having the angles θ_1 and θ_2 for each face of the polygon, the differences between adjacent faces are computed as

$$D(N) = \theta_N - \theta_{N+1},$$

where $N = 1, \dots, 12$, and $\theta_{13} = \theta_1$. Then, the error of each face referred to the angle that each face should have is given by:

$$E(N) = D(N) - \frac{1}{12} \sum_{N=1}^{12} D(N) \quad (5)$$

3. EXPERIMENTAL RESULTS

The experimental set up depicted in Fig. 1 was build. In order to minimize the power variations of the beams, an intensity stabilized He-Ne laser was used (Spectra Physics Mod. 117A); its beam has also better directional stability. A Melles-Griot lens 01 LAO 357, was used; it has a focal length of 800 mm and a diameter of 63 mm. A laser picowatt digital power meter from Newport, mod. 835, together with a PIN Silicon photodetector mod. 835UV, was used to sense the power variations. The knife edge was build with a piece of Aluminium sharpened on the edge, and it was mounted on the detector case. The output power of the laser was of 1.9 mW, but after being divided and reflected the power of each beam at the focal plane, was reduced to around $P_{T1} = 0.6$ mW and $P_{T2} = 0.4$ mW; these values were changing a small amount with the face where the beam was reflected.

The radius of each beam at the knife edge position was measured by the opaque wire method, as was explained above; five wires of different diameters, calibrated with an accuracy of 0.0001 “ were used. The radius of each beam was: $r_{o1} = 467 \pm 2 \mu\text{m}$ and $r_{o2} = 481 \pm 6 \mu\text{m}$.

In Fig. 2 the error for each face of the polygon, measured with the laser, is shown.

4. DISCUSSION AND CONCLUSIONS

From Fig. 2, it can be appreciated the value of the proposed method. The error bars represent plus-minus one standard deviation after three different measurements of the errors of the faces of an optical polygon. The highest standard deviation for the faces 90 and 180 degrees are less than one and a half arc seconds. This is a fine value to calibrate such components. Perhaps the main disadvantage of the method, lies on the actual experimental set up; presently, aligning the set up and making manually all the measurements, take several hours; the method, however, is suitable for being automated by using mechanical actuators to smoothly rotate the polygon and by

acquiring the laser power data with the help of a PC computer. The experimental results here reported show the capabilities of the method.

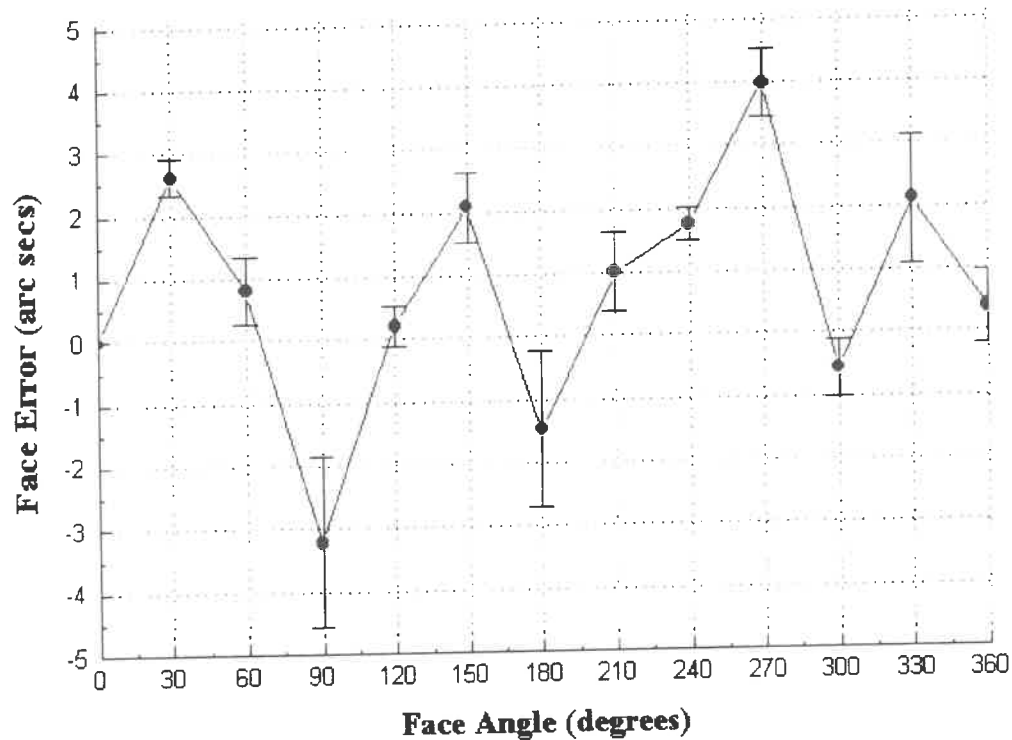


Figure 2. Experimental results for the calibration of an optical polygon with a laser beam.

5. REFERENCES

- [1] K.J. Hume, *Metrology With Autocollimators*. Hilger & Watts Ltd. (London, England, 1965).
- [2] R. Nava, *Calibración de un Polígono Óptico de 12 Caras Mediante Dos Autocolimadores (Calibration of a Twelve Faces Optical Polygon with Two Autocollimators)*. Technical Report CIUNAM, (México, 1992). In Spanish.
- [3] Díaz Uribe, R., Rosete Aguilar, M., Ortega Martínez, R., "Position Sensing of a Gaussian Beam with a Power Meter and a Knife Edge", *Rev. Mex. Fís.*, **39** (1993) 484-492.
- [4] Rufino Díaz-Uribe, Roberto Ortega-Martínez, "Laser Beam Centering by the Opaque Wire Method", in *Second Iberoamerican Meeting of Optics*, Daniel Malacara, Sofia Acosta, Ramón Rodríguez Vera, Zacarias Malacara, Arquímedes Morales, Editors, Proc. SPIE 2730, 576-578 (1996).

6. ACKNOWLEDGMENTS

We acknowledge the valuable help of Miss Fanny Pineda Pineda in the measuring process of this work.

Nano-Metrology of Roundness and Spindle Error

Wei Gao, Satoshi Kiyono and Takamitsu Sugawara
Department of Mechatronics, Faculty of Engineering
Tohoku University, Sendai, 980 JAPAN

ABSTRACT

This paper presents a new measurement system for nano-metrology of rotational errors of air spindles and roundness errors of master balls. The system is based on the orthogonal mixed method, which can separate the roundness error and the rotational error with the accuracy of the probes. An optical sensor consisting of one displacement probe and one angle probe with an angular distance of 90° was constructed. Both the displacement probe and the angle probe utilize the principle of critical angle method of total reflection, and has the resolution better than 1nm and 0.01 aresec, respectively. The measurement results show that the roundness measurement can be performed with a repeatability of several nanometers.

Key Words : metrology, roundness error, spindle error, sensor, software datum, error separation

INTRODUCTION

Nano-metrology of roundness errors of cylindrical workpieces and spindle errors of machine tools is becoming increasingly important. In order to identify these errors individually, it is necessary to separate the roundness error and spindle error from each other in the accuracy of nanometer level. Methods using multiple probes are the most widely used error separation method [1]. Such methods are also suitable for on-machine measurements because the repeatability of the spindle error is not necessary. However, only displacement probes are used in the traditional multi-probe methods, some high-frequency components cannot be accurately measured. This weakness limits the application of such methods to nano-metrology.

We have proposed a new multi-probe method which is named as the mixed method [2]. This method utilizes not only the information of the displacement probe but also that of the angle probe, and can separate the roundness error and the spindle error from each other with the accuracy of the probes.

In this paper, we propose an optimum mixed method which is named as the orthogonal mixed method. The orthogonal mixed method uses one displacement probe and one angle probe with the angular distance of 90° and has the best characteristics. It is evident that nano-metrology of roundness errors and spindle errors can be realized if we use the probes with nano-accuracy to construct the

orthogonal mixed method. We also presents a new system for nano-metrology of rotational errors of air spindles and roundness errors of master balls.

PRINCIPLE OF ERROR SEPARATION

Figure 1 shows the principle of the orthogonal method for separating the roundness error and the spindle error schematically. Two probes are fixed around a cylindrical workpiece. Let P be a representative point of the workpiece and the roundness error be described by the function $r(\theta)$, where θ is the angle between point P and the X axis. The output $m_A(\theta)$ of the displacement probe A and the output $\mu_B(\theta)$ of the angle probe B can be expressed as follows:

$$m_A(\theta) = r(\theta) + e_X(\theta), \quad (1)$$

$$\mu_B(\theta) = r'(\theta - \pi/2) + e_X(\theta)/R_r. \quad (2)$$

Here, $e_X(\theta)$ is the X-directional component of the spindle error, and R_r is the radius of the workpiece.

Therefore, the differential output $m_{om}(\theta)$, in which the roundness error is separated from the spindle error, can be denoted as

$$\begin{aligned} m_{om}(\theta) &= m_A(\theta) - R_r \mu_B(\theta) \\ &= r(\theta) - R_r r'(\theta - \pi/2). \end{aligned} \quad (3)$$

The transfer function $H_{om}(\omega)$ to determine the relation between $r(\theta)$ and $m_{om}(\theta)$ can be defined as

$$\begin{aligned} H_{om}(\omega) &= M_{om}(\omega)/R(\omega) \\ &= 1 - j\omega e^{-j\omega\pi/2}, \end{aligned} \quad (4)$$

where ω is the spatial frequency, and $M_{om}(\omega)$ and $R(\omega)$ are the Fourier transforms of $m_{om}(\theta)$ and $r(\theta)$, respectively. As shown in Fig. 2, the amplitude of $H_{om}(\omega)$ does not become zero in the whole frequency range when $\omega > 1$ just like the general mixed method [2]. $R(\omega)$ can be obtained from $M_{om}(\omega)$ and $H_{om}(\omega)$, and $r(\theta)$ can be evaluated from IFFT of $R(\omega)$. The spindle error $e_X(\theta)$ can also be calculated by using $r(\theta)$ and the probe output.

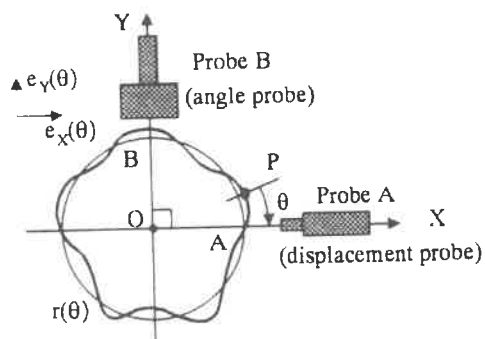


Fig. 1 Probe setup of the orthogonal mixed method

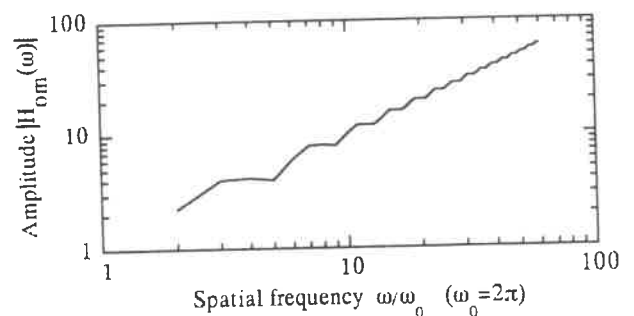


Fig. 2 Transfer function of the orthogonal mixed method

EXPERIMENT

Figure 3 shows the experimental setup for roundness measurements by using the orthogonal mixed method. An optical sensor consisting of one displacement probe and one angle probe with an angular distance of 90° was constructed. To measure the roundness error and the spindle error with nanometric accuracy, a displacement probe with a resolution of 1 nm and an angle probe with a resolution of 0.01 arcsec are necessary. Here we adopted the displacement probe and the angle probe utilizing the principle of critical angle method of total reflection [4]. A master ball with a diameter of 1 inch which is mounted on an air spindle is used as the target.

Figure 4 shows the results of two repeated calibrations of the displacement meter. A capacitance type displacement probe (Microsense) was used as the reference. The fitting errors to a polynomial approximation of degree 3 were less 0.5% of the calibration range. A photoelectric autocollimator was used as the reference angle meter to calibrate the angle probe. Two repeated calibration results were plotted in Fig. 5. The fitting errors to a 3 degree polynomial approximation of degree 3 were also less than 0.5% of the calibration range.

Stability tests of the optical sensor showed that the displacement probe and the angle probe have the stability of 1 nm and 0.01 arcsec, respectively.

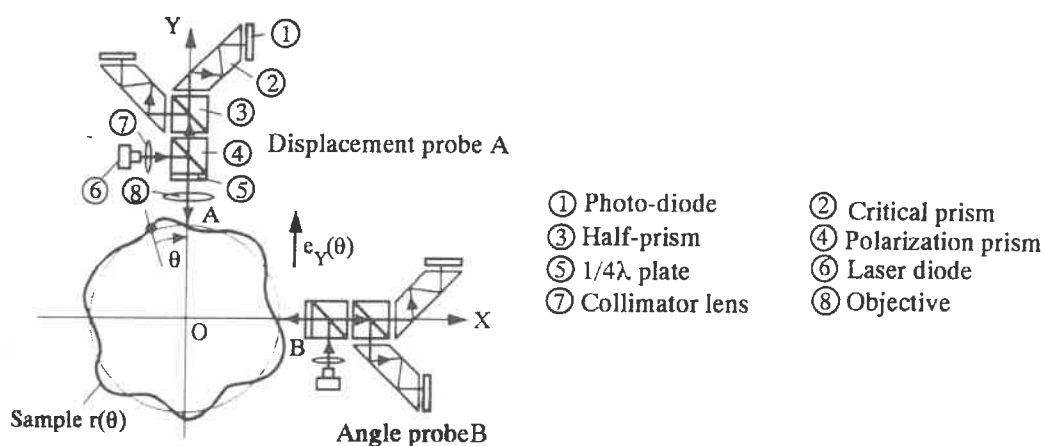


Fig. 3 Optical sensor for orthogonal mixed method

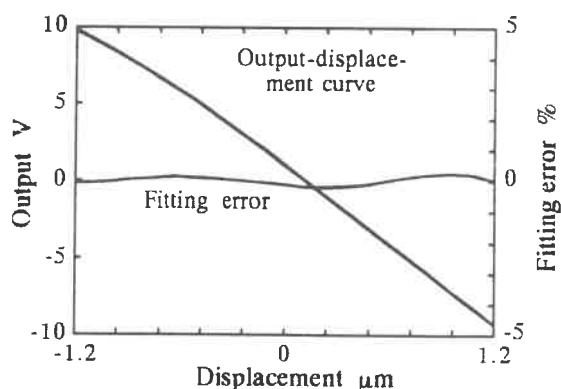


Fig.4 Calibrated results of the displacement probe

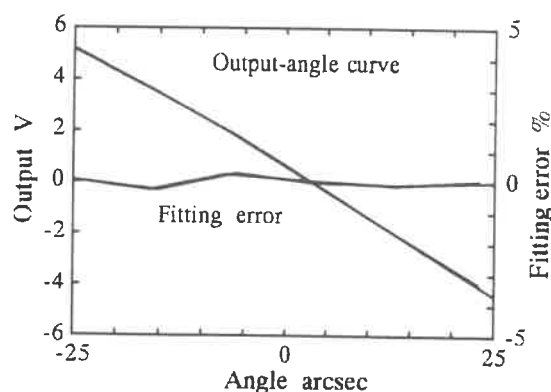


Fig.5 Calibrated results of the angle probe

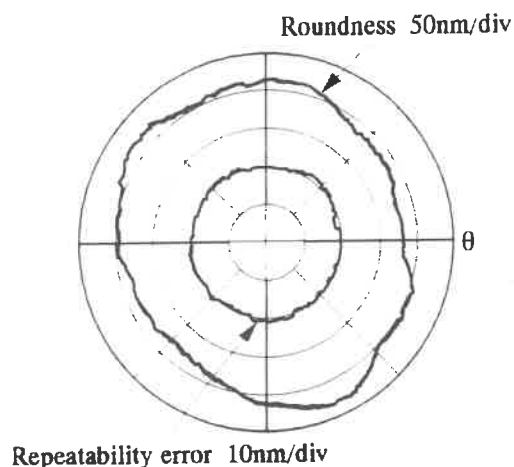


Fig. 6 Roundness errors $r(\theta)$ and the repeatability error $\Delta r(\theta)$

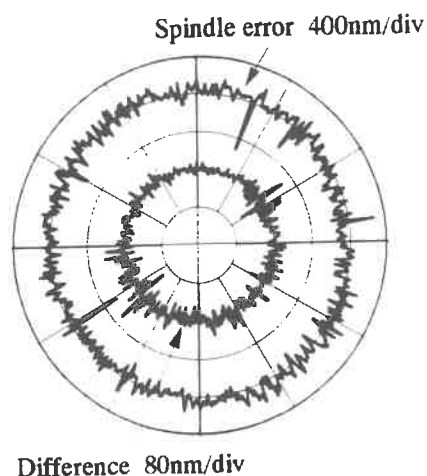


Fig. 7 Spindle errors $e_v(\theta)$ and the difference $\Delta E_v(\theta)$

Figure 6 shows the measured roundness errors of two repeated measurements and the repeatability error between the two measured results. The sampling number was 512. It can be found that the roundness error is about ± 30 nm, and the repeatability error is about ± 2 nm. The repeatability error might be caused by the positioning error of sampling. The measured spindle errors of the two repeated measurements and the difference between them are shown in Fig. 7. The spindle error is about ± 400 nm, and the difference is about ± 70 nm.

CONCLUSIONS

- 1) A new error separation method for roundness measurements has been proposed.
- 2) An optical sensor has been constructed to realize the orthogonal mixed method in nanometric accuracy.

ACKNOWLEDGMENT

The authors would like to thank Mr. S. Fukushima of Canon Ltd and Micro Surface Engr. Inc. for providing the air spindle and the master ball, respectively. Financial support was provided by the Mechatronics Scientific Foundation and Grant-in-Aid for Scientific Research of The Ministry of Education of Japan.

REFERENCES

1. D. J. Whitehouse: Some theoretical aspects of error separation techniques in surface metrology, *Journal of Physics E: Scientific Instruments*, Vol. 9, (1976), pp.531-536.
2. Wei Gao, Satoshi Kiyono and Tadatashi Nomura: New Multi-probe Method of Roundness Measurements, *Precision Engineering, Journal of the American Society for Precision Engineering*, 19-1, (1996), pp. 37-45.
3. T. Kohno, N. Ozawa, K. Miyamoto, and T. Musha: High precision optical surface sensor, *Applied Optics*, 27-1, (1988), pp. 103-108.
4. P. Huang and S. Kiyono: Angle measurement based on the internal-reflection effect: A new method, *Applied Optics*, 31-28, (1992), pp. 6047-6055.

SPINDLELESS INSTRUMENT FOR THE ROUNDNESS MEASUREMENT OF PRECISION BALLS

Eugene Gleason, Heinrich Schwenke
Bal-Tec, Division of Micro Surface Engr.

Abstract

This paper presents a realized measuring instrument for the roundness measurement of balls. Instead of employing a spindle, it uses the surface of the ball itself as a geometrical reference. This instrument enables us to measure the roundness of one inch steel balls with a nanometer resolution.

Motivation

Mechanical mechanisms have a high demand for rotational elements: Especially for ball bearings high quality balls are a necessary prerequisite for accurate and quiet operation and a long duty time. But also for kinematic components and standards for dimensional metrology balls with high precision demands are needed. Bal-tec Engr. manufactures routinely balls to a guaranteed roundness of 100 nm. This arrives at the limit of conventional measuring techniques. To further expel the accuracy of manufacturing, the measuring capabilities had to be improved significantly.

Conventional roundness measuring instruments employ a spindle which serves as a reference. Consequently any radial error of the spindle motion appears as an error of the measurement, which results in the necessity of a spindle accuracy that is five to ten times higher than what one wants to measure. For a desired measuring uncertainty of 5 nanometers there is no such reference motion available. This led us to the solution of using the measured object itself as a reference, which is possible since we only measure differential values.

Principle

The principle is based on the three point measurement method, that is discussed in various publications [1,2,3]: Figure 1 shows the arrangement in principle.

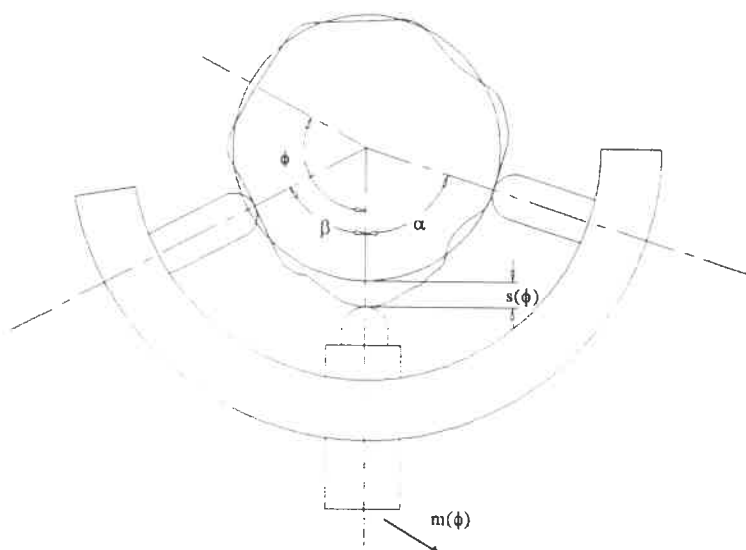


Figure 1: Arrangement in principle

The cross-section of the ball (or a shaft) is contacted at three points, one of them serves as a transducer. An out of roundness results in a change of the transducer displacement depending on the angular

position of the workpiece. This signal unfortunately cannot directly serve as a value for the form error at the position of the transducer, since the form at the two other points effect the signal, too. The functional relationship can be established as [4]:

$$m(\phi) = s(\phi) - \frac{\sin(\alpha)}{\sin(\alpha + \beta)} s(\phi - \beta) - \frac{\sin(\beta)}{\sin(\alpha + \beta)} s(\phi + \alpha)$$

The reconstruction of the profile can be accomplished by Fourier transformation into a set of n harmonics [4]. It shows up the sensitivities of the single harmonics strongly depend on their order and the angles α and β .

Therefore a harmonic correction must be applied before re-transformation. But also these angles can be used to optimize the sensitivity for special harmonics. We chose $\alpha = \beta = 60^\circ$, which results in a high amplifications for the critical low order harmonics (due to the manufacturing process these shapes tend to occur most significantly). Figure 2 shows the amplification spectra for two combinations of α and β .

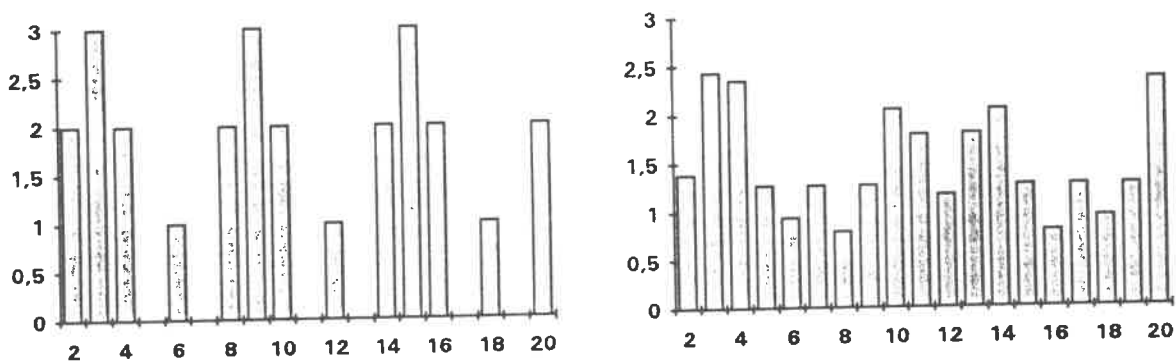


Figure 2: Spectra for two combinations of α and β

The Transducer

The high sensitivity, its simple mechanical design and its relatively low cost let us choose a capacitive transducer as a gage. The gage is measuring directly against the surface of the ball, which itself serves as one capacitor plate. By that means, any moving mechanical parts is avoided. We apply corrections for the effect resulting from the spherical surface. The system consisting of the capacitive gage and a driver unit proved to have the sensitivity of 4 mV /nm stated by the manufacturer of the system. It also showed up to have a very low noise if mounted in a mechanical stiff set up: Employing relatively long sampling intervals (0.5 s) a signal stability of one nanometer during one minute was accomplished.

Instrument Design

The desired accuracy is accomplished by a simple and stiff design of the instrument. The two contact points in the measuring plane are realized by tungsten carbide cylinders with diamond tips. These tips serve as the actual contact points to the surface of the ball and have an extremely elaborate surface finish with a roughness of a few angstrom (0.1 nm). The contact point in the back -perpendicular to the measuring plane- is as a brazed gold contact to assure good electrical conductivity for the capacitor. The sensor itself is clamped in a support, which can be adjusted in two directions by elastic deformation of flex joints. These flex joints are integrated in the support of the sensor by EDM machining to avoid any stick and slip effects during the adjustment. For the material of the instrument base we chose aluminum, because of its high thermal conductivity and its good machineability. Figure 3 shows the base of the instrument and the support, naming the functional components of the device.

Automating the Measuring Process

The instrument has been designed to allow the fast and reliable check of high precision balls after the manufacturing. For this purpose a PC controlled manipulator was designed, that is able to rotate the ball about two axes. For the roundness measurement of a single cross-section the manipulator lifts the ball, turned it about a certain angle and put the ball back in the measuring position. After a short time (0.5s) the sensor signal is sampled and the manipulator turns the ball again. This procedure is carried out 32 to 128 times during one measurement. This is equivalent to an angular resolution of 11.25° to 2.8° . The second axis of the manipulator is used to rotate the ball in another measuring plane to perform fully automated measurements of three perpendicular cross-sections. The manipulator and the read out of the sensor signal is controlled by a PC that also evaluates the results by Fourier analysis and visualizes the results. The prototype is able to make a automated complete check of a ball in 15 minutes.

Environmental Effects

Measuring in this almost atomic range, careful environmental insulation is a important prerequisite. As the dominant environmental effects we considered temperature, noise, vibration and electromagnetic interference. Especially the awareness of temperature changes during the time of measurement (ca. 4 minutes for one cross-section) led us to a "box in the box" design like it is shown in figure 4. We set up an enclosure of 1/8 inch thick copper plates (3) around the actual measuring instrument (5). These plates serve to shield thermal radiation and as a heat sink they reduce thermal gradients. In addition, they shield electromagnetic interference. The inner box is based on a strong wooden table, which is vibration insulated by static air bearings (6). The manipulator (4) to lift and turn the ball is placed outside the copper box. The whole arrangement is enclosed by a foam insulated steel box (1,2).

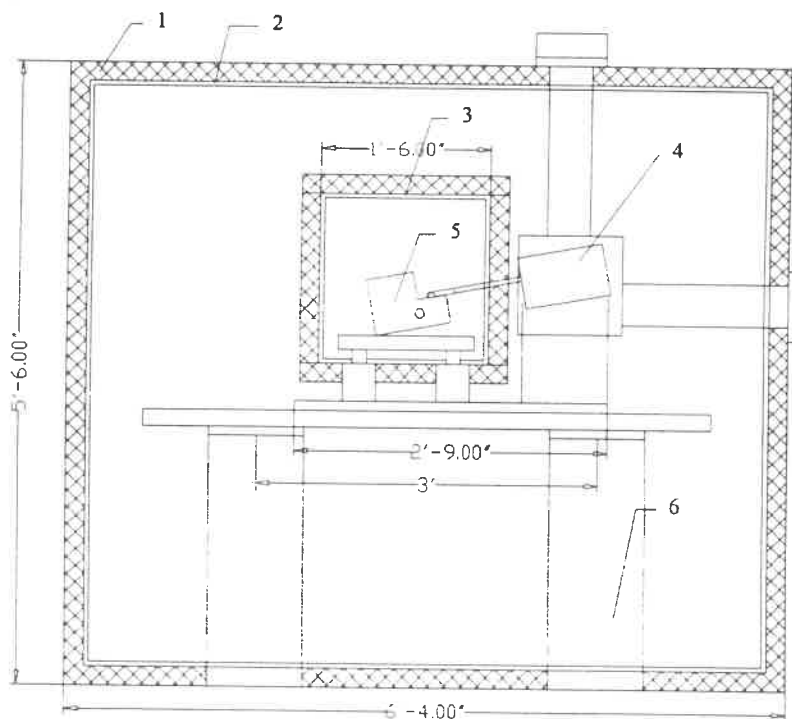


Figure 4: The "box in the box" design for environmental insulation

The temperature related changes of the signal is reduced to less than two nanometers over five minutes. The "box in the box" design also reduces the influence of vibration, noise and electromagnetic interference to a negligible level.

Results

We measured one inch balls with the quality grade #25 (out of roundness less than 25 microinches = 635 nm) up to quality grade #3 (< 3 microinches = 76.2 nm). The repeatability of the measurements was within 10 %. Figure 5 shows two measurements performed on two different days on the same ball. For this ball we measured an out of roundness of 40 nm with a repeatability of 2 nm.

For comparative measurements with a conventional instrument balls with larger form errors had to be employed, since the resolution of the conventional instrument was limited to a resolution of ca. 30 nm. The agreement of the results of grade #25 balls was within this resolution, even when performed by different operators.

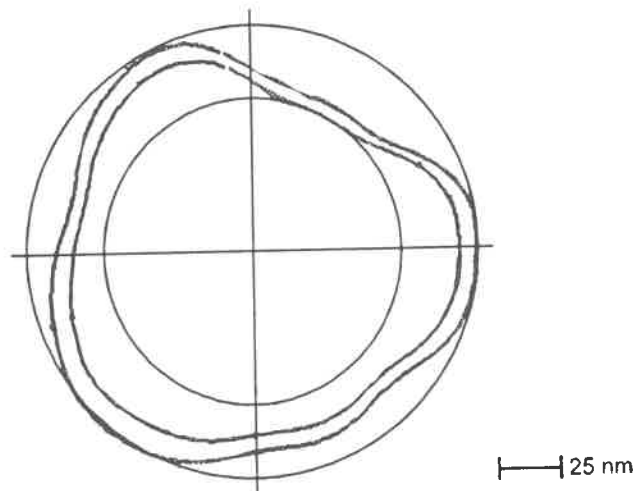


Figure 5: Comparison of two independent measurements within two days

Conclusion

Already the prototype proved the capability of measuring the roundness of balls as accurate as the worlds most advanced measuring instruments. We measured balls of less than 50 nm out of roundness with a repeatability of 2 nm. In general, the repeatability of the measurements is within 10% and the agreement with a conventional spindle instrument is within its resolution of 30 nm. The PC controlled manipulator performs fully automated measurements of the balls in three perpendicular planes in 15 minutes. Software to control, evaluate and visualize the measurement has been developed. Limiting factors of the present design are the "blindness" for certain harmonic orders and the restriction to balls with a conductive surface and a certain size. For Microsurface Engr. the capability of measuring balls to a nanometer order is a step to further improve the quality of the high precision balls.

Bibliography

- [1] Barth "Kombinierte Zweipunkt-Dreipunkt-Messung bei gleichzeitigem Bestimmen der Kreisformabweichung", Feingerätemeßtechnik 33 (1984), Springer Verlag, Berlin 1984, Germany
- [2] Moore "Design Considerations in Multiprobe Roundness Measurement", Journal of Physics E: Scientific Instruments 1989 Volume 22, Great Britain
- [3] Shinno, Mitsui and Tatsue "A New Model for Evaluating Error Motion of Ultra Precision Spindles", Annals of the CIRP Vol. 36/1/1987

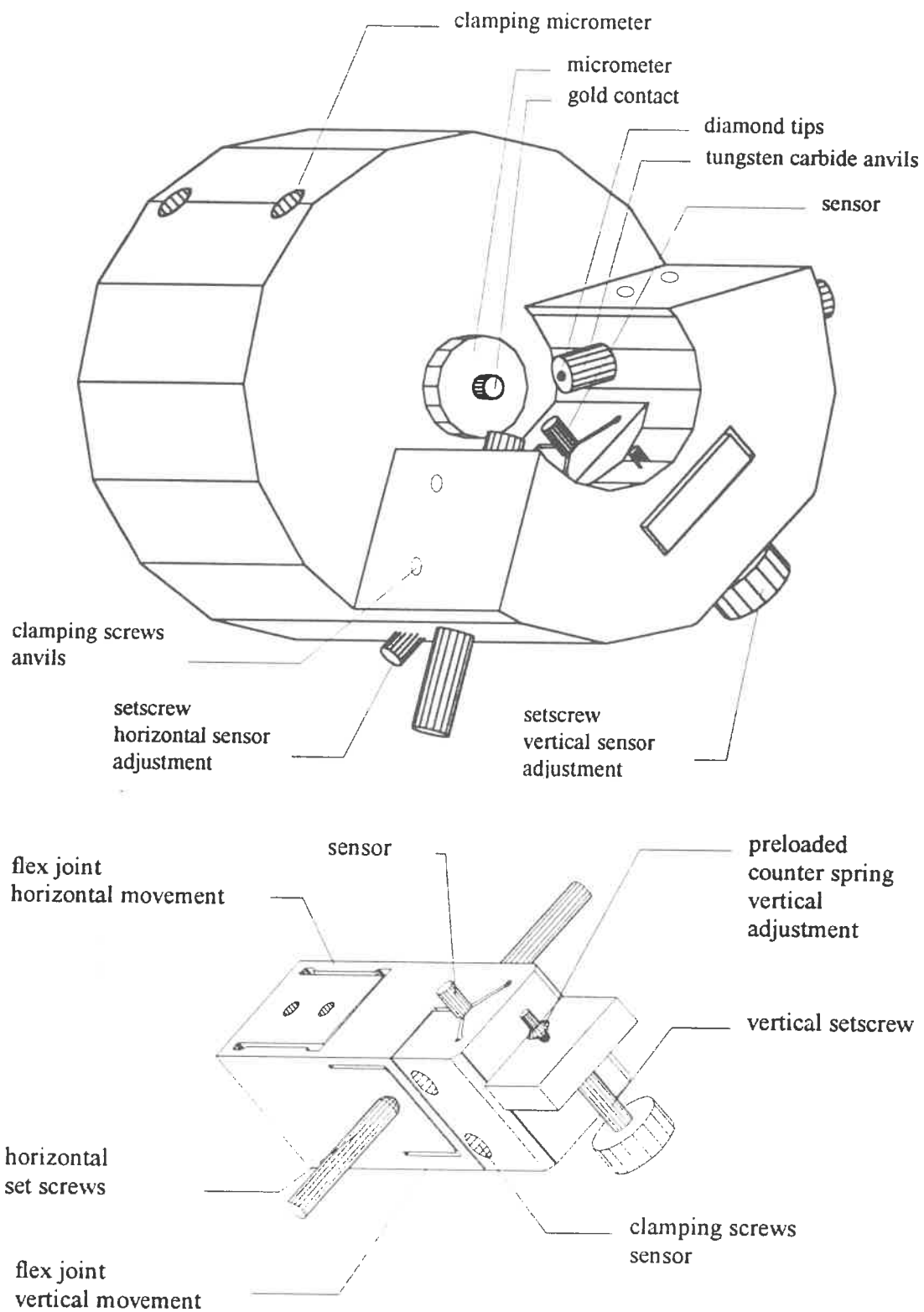


Figure 3: Sketch of the instrument base and the sensor support

ASSESSMENT OF DRILL POINT GEOMETRY BY GENETIC ALGORITHM AND STEEPEST GRADIENT METHOD

Lakshman HAZRA*, Ryohei NAKAMURA*, Takaharu KURODA**, Hideo KATO*,
Yoshihiro TSUCHIYA† and Itaru SAKUMA†

* Chiba University, ** Kisarazu National College of Technology, † Fujita Works, JAPAN

1. Introduction

Drill is one of the most basic cutting tools in machining process. The geometry of drill point affects productivity and accuracy of the hole making process and the processes following. The point of worn twist drill, or the flank surfaces, are reground to sharpen the cutting edges by using drill point grinder. However it is difficult to shape the surface into the target geometry precisely, because of inevitable positioning error, wear of grinding wheel and so on. Accordingly, in order to maintain the cutting performance of the drill, the inspection of the geometry is very important. But conventional drill point geometry inspection methods using metrological instruments, such as simple gauges, profile projector and tool maker's microscope, call for operator's skill and time consumption¹⁾. Ramamoorthy et. al.²⁾ have developed a method by using image processing via noncontact coordinate measuring machine which is simple but it is only for some representative values such as lip clearance angle and chisel edge angle. This paper deals with a new evaluation method for complete geometry of conically ground drill point, aiming to develop a computer aided inspection system (CAI) of it. In this method, from the kinematic relationship between the drill point and grinding wheel when regrounding, five geometrical parameters have been developed for evaluation.

2. Geometrical Parameters for Drill Point

Figure 1 shows the relative position between grinding wheel and drill point. The X-Y-Z system's origin is the center of chisel edge of drill, and Z-axis identifies with drill axis. The x-y-z system's origin is the top of the grinding cone, z-axis identifies with grinding cone axis. Denoting the drill point angle by α and the drill diameter by D , x-y-z system equals X-Y-Z system which turns ζ around Z-axis, translates aD along new X-axis, eD along new Y-axis and pD along new Z-axis, and turns δ around new X-axis. Therefore, coordinates (X_i, Y_i, Z_i) are transformed into coordinates (x_i, y_i, z_i) by Eqs.(1) to (5). In x-y-z system, the ideal cone is expressed as Eq.(6). When measuring the coordinates of n points on the flank surface in X-Y-Z system (Fig.2),

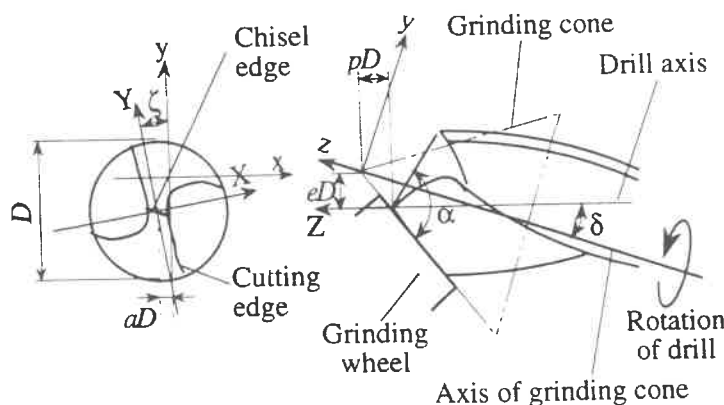


Fig.1 Geometrical relationship between drill and grinding wheel when regrounding

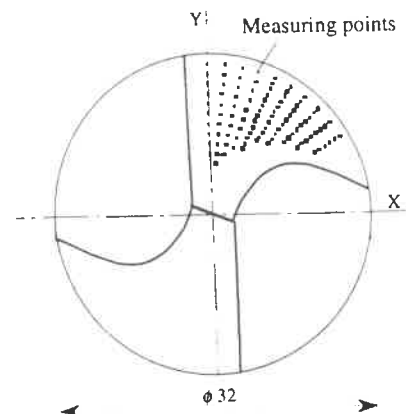


Fig.2 Measuring points on flank surface

$$\begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\delta & -\sin\delta & 0 \\ 0 & \sin\delta & \cos\delta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -aD \\ 0 & 1 & 0 & -eD \\ 0 & 0 & 1 & -pD \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\zeta & -\sin\zeta & 0 & 0 \\ \sin\zeta & \cos\zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} \quad (1)$$

$$p = (-k_2 - \sqrt{k_2^2 - k_1 \cdot k_3}) / k_1 \quad (2)$$

$$k_1 = \sin^2\delta - \cos^2\delta \cdot \tan^2(\alpha/2 - \delta) \quad (3)$$

$$k_2 = -e \cdot \cos\delta \cdot \sin\delta \cdot \{1 + \tan^2(\alpha/2 - \delta)\} \quad (4)$$

$$k_3 = a^2 + e^2 \cdot \{\cos^2\delta - \sin^2\delta \cdot \tan^2(\alpha/2 - \delta)\} \quad (5)$$

$$x^2 + y^2 = z^2 \cdot \tan^2(\alpha/2 - \delta) \quad (6)$$

$$f(a, e, \zeta, \delta, \alpha) = \frac{1}{n} \sum_{i=0}^n \left(z_i - \frac{\sqrt{x_i^2 + y_i^2}}{\tan(\alpha/2 - \delta)} \right)^2 \quad (7)$$

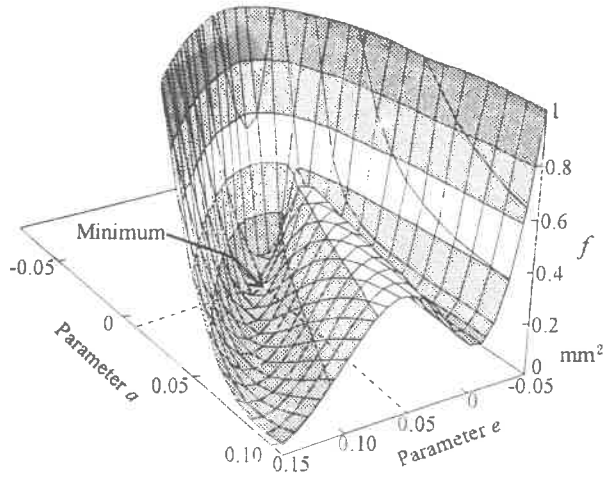


Fig.3 Shape of Index f around its minimum

or (X_i, Y_i, Z_i) , $i=1, \dots, n$, minimization of Index f given by Eq. (7), which is the mean square of relative error in z -axis direction between the coordinates and the ideal cone, estimates the parameters $(a, e, \zeta, \delta, \alpha)$. And the parameters of both flank surfaces reasonably describe the drill point geometry.

3. Evaluation of Parameters

Index f is nonlinear in terms of the parameters and looks multimodal, as shown in Fig.3, which is obtained from numerical simulation. Firstly, we examined Steepest Gradient Method to this minimization problem by using simulation data instead of measured coordinates. In Fig. 4, the change of parameter α against parameter

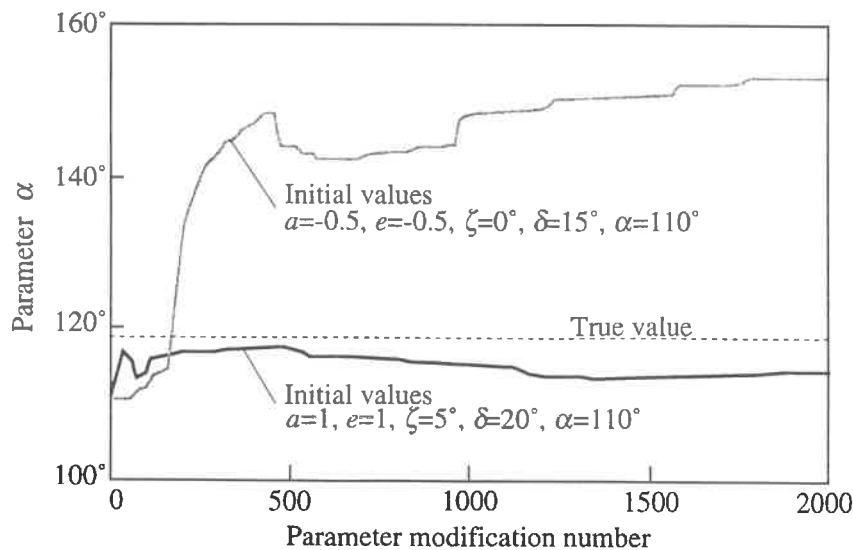


Fig.4 Change of parameter α by steepest gradient method

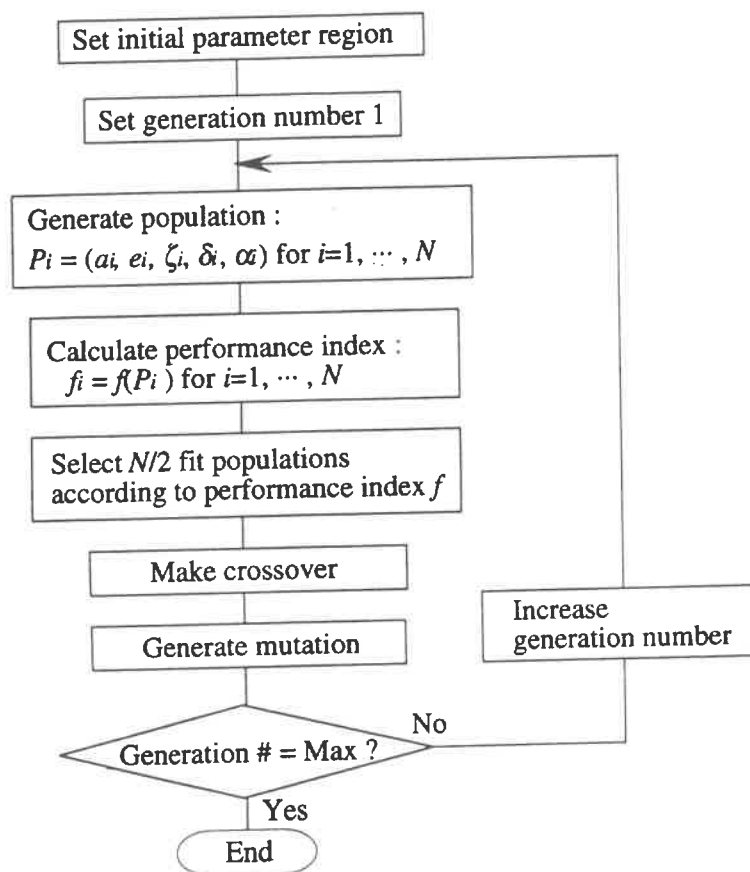


Fig.5 Flow chart of genetic algorithm

modification number is shown in the cases of two different sets of initial parameter values. As shown in this figure, Steepest Gradient Method sometimes makes the parameter values go far away from true value, unless the initial values are near the true values. Meanwhile, it is difficult to find proper initial parameter values in general. Then, we examined Genetic Algorithm(GA). The flowchart of GA is shown in Fig.5. Population number $N=250$ and mutation rate 10% are adopted on the basis of preliminary numerical experiments. Change of parameter α by GA against generation number is shown in Fig.6. In this case, the initial parameter values should be given as the ranges, or $a=-1\cdots 1$, $e=-1\cdots 1$, $\zeta=-10^\circ\cdots 10^\circ$, $\delta=10^\circ\cdots 20^\circ$, $\alpha=100^\circ\cdots 120^\circ$. It is found in the figure that GA makes the

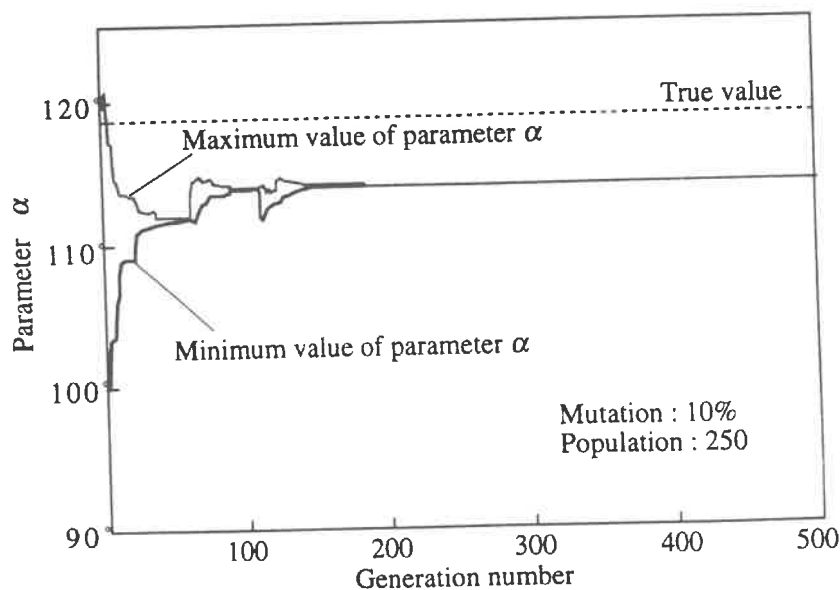


Fig.6 Change of parameter α by genetic algorithm

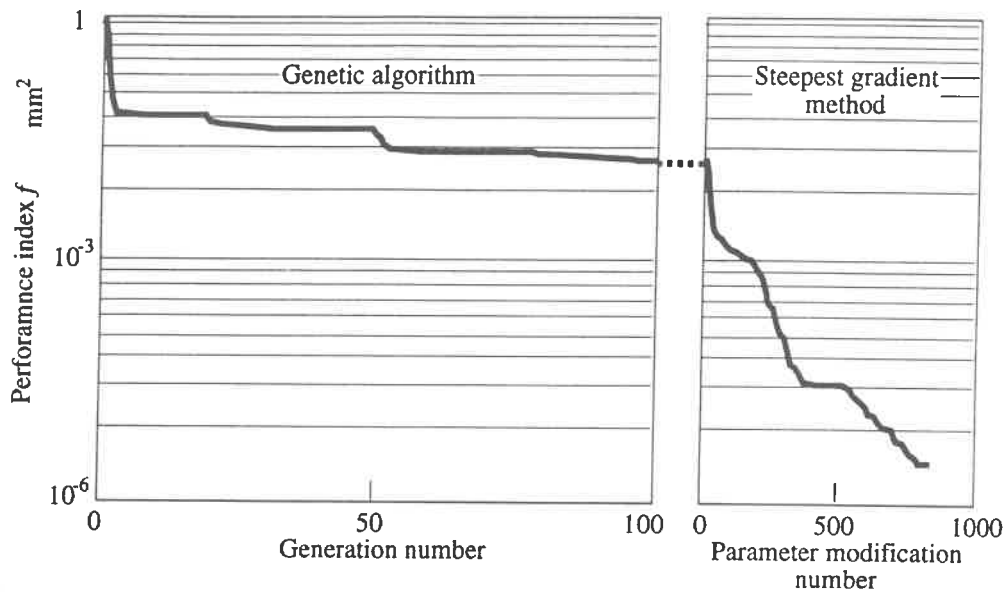


Fig.7 Change of performance index f by genetic algorithm and steepest gradient method (Simulation)

Table 1 Comparison between true and estimated values

Parameter	a	e	ζ	δ	α	f
Initial region	-1~1	-1~1	-10°~10°	10°~20 °	100°~120°	—
True value	0.01	0.05	5°	20°	118°	0
Estimated value	0.01131	0.05247	5.89°	20.03°	117.418°	$2.6 \times 10^{-6} \text{ mm}^2$

parameter α converge, but the value hardly improves after about 200th generation. So, we tried Steepest Gradient Method after applying GA to some extent. In Fig.7, the change of Index f by using GA until the 100th generation and subsequently Steepest Gradient Method is shown. As can be seen in this figure, Index f reduces down drastically after change of method. The parameter values estimated by this method are shown in Table 1 together with the true values. The estimation accuracy depends on the condition for judging convergence. Calculation time required is 3min49s for GA and 38s for Steepest Gradient Method by using a PC(Pentium120MHz, OS:Windows95, Compiler:Turbo C++4.0).

4. Conclusions

- (1) By analysing the relative motion of the grinding wheel and the drill point, a method to assess the point geometry is developed.
- (2) The results show that the combined method of GA with Steepest Gradient Method is suitable for the identification of the parameters from the viewpoints of calculation speed and accuracy.

References

- 1) D. B. Dallas : Tool and Manufacturing Engineers Handbook -- 3 rd. Edition, McGraw-Hill (1976) pp 3-3- --3-23.
- 2) B. Ramamoorthy and V. Radhakrishnan : Computer Aided Inspection of Cutting Tool Geometry. Precision Engineering, 14, 1 (1992) pp 28-35.

Transmission Electron Microscopy of Precision Machined Surfaces

Toshiaki Kaneeda, * Seiichi Yokomizo** and Daisuke Higashi *

* : Department of Mechanical Engineering, Okayama University of Science
1-1 Ridai-cho, Okayama, 700 JAPAN

** : Industrial Technology Center of Okayama Prefecture
5301 Haga, Okayama, 701-12 JAPAN

Abstract

Transmission electron microscopy has been conducted about precision and ultra precision machined surfaces from a metal physical view point. Transmission electron microscopy showed that tangled dislocation networks were dynamic-recovered to produce a number of subgrains in all but the ultra precision cuttings. No significant differences in the microstructure between conventional and precision machinings could be recognized in spite of different tool material and cutting conditions. However, most of the microstructure in the ultra precision machined surfaces developed into not subgrain but cell structure.

Key words : Machined surface, transmission electron microscope, precision machining, ultra precision machining, microstructure, subgrain, cell, texture, deformed layer.

1. Introduction

Examination of machined surface can be usually divided into two groups. One is the external characteristics, surface roughness, waviness, and so on. The other is invisible characteristics, that is the deformed layer, residual stress, hardness, etc. Precision cuttings have recently become a part of conventional manufacturing process, moreover, ultra precision cuttings may have been accepted as a kind of precision machining. The examinations mentioned above are not sufficiently effective for ultra precision machined surfaces. Therefore, internal microstructure of precision and ultra precision machined surfaces should be critically examined with an appropriate examination. Kaneeda et al.¹⁾ has already revealed microstructure of the machined surface, chip, shear plane and splitting point in pure Al workmaterials, however, that was for conventional machining. This study actually focuses on precision and ultra precision cuttings.

Transmission electron microscopy was carried out on machined surface layers of pure Al and Al-Mg alloy. It is the purpose of this paper to create a better understanding of microstructure, in particular, dislocated internal structure of precision and ultra precision machined surface layers using a high accelerating

voltage, 200kV transmission electron microscope (TEM).

2. Experimental procedure

2.1 Precision cutting

2.1.1 Precision cutting experiments

Precision cutting experiments were conducted on a NC orthogonal precision cutting equipment with cutting speed $V=25.7$ m/min. Minimum depth of cut was $t_1=1\mu\text{m}$. Pure Al (A1050-0, more than 99.5% purity) was used as the workmaterial. Diamond tools having 0° rake angle were selected. Ethyl alcohol were used as a cutting fluid.

A quartz force transducer (Kistler Co. type 9251A), having high sensitivity, stiffness, and natural frequency was employed to measure the cutting forces. The transducer offers $1000\text{N}/\mu\text{m}$ (tangential force

Table 1 Precision cutting conditions

Work material	Pure aluminum (A1050-0)
Dimensions of work	$3.40 \times 3.5 \times 3$
Cutting form	Orthogonal
Tool material	Diamond
Rake angle α	0
Relief angle ϵ	10
Depth of cut t_1 μm	1
Cutting fluid	Ethyl alcohol

Table 2 Ultra precision cutting conditions

Work material		Al-Mg alloy
Dimensions of work		$\phi 65 \times 35$
Tool		Single point diamond tool
Cutting form		Surfacing
Tool angle	Rake angle α	0°
	Relief angle ϵ	7°
Nose radius mm		1
Tool edge angle θ		90°
Main spindle speed mm/min		0 – 408.4
Feed rate μ m/rev		2
Depth of cut t_1 μ m		2, 4 & 8
Cutting fluid		White kerosene (Mist type)

direction) and 300N/ μm (normal force direction) in stiffness, and 17.678 and 9.683 kHz in natural frequency respectively.

The cutting speed was a constant 25.7m/min, as microvicker's hardness distribution in the machined surface by the cemented carbide tool cutting has already been revealed by the author at this cutting speed.²⁾

Table 1 summarizes the precision cutting conditions.

2.1.2 Cutting methods

At first, the workmaterial sectioned from the Al plate was annealed at 400° C for 2 hours. The workmaterial annealed was carefully machined in order to obtain uniform depth of cut in the cutting experiments. This process should not produce a large deformed layer of the machined surface, therefore, 1 μm depth of cut was selected as a pre-cutting condition.

2.2 Ultra precision cutting

Ultra precision cutting experiments were conducted on a Rank Pneumo Nanoform 600 to obtain ultra smooth mirror surfaces like hard disk substrates. The Al-Mg alloy disk having a 65mm diameter for a hard disk substrate was used as the workmaterial. Diamond tools had 0° rake angle and 10° clearance angle. Main spindle turning speed was constant 2000rpm, so the

cutting speeds ranged from 62.8m/min to 408.2m/min. Feed rate was 2 $\mu\text{m/rev}$. Ethyl alcohol were used as a cutting fluid. Before the cutting experiments, 2 hour - 280° C-annealing was carried out to remove the work hardened layer.

Table 2 summarizes ultra precision cutting conditions.

2.3 Specimen preparation for observation

Fig. 1 shows the observation point of the specimen picked from the workmaterial machined. It is very important to obtain thin films without any damage during processing. First, the workmaterial was sliced with less than 0.2mm thickness by an electric discharge machine with a wire cutting device along the cutting direction. The thin plate was then electro-chemical-jet polished by a Tenupole-3 (Storeurs Co.) to make a concave lens-shaped thin film. Less than 1 μm thickness Al film can be penetrated by an electron beam with 200 kV accelerating voltage.

3. Experimental results and discussion

3.1 Machined surface properties

Table 2 shows the machined surface properties. The results indicated that smaller depth of cut offered smaller cutting forces, smoother surfaces and less hardness values as usually expected in the precision machining.

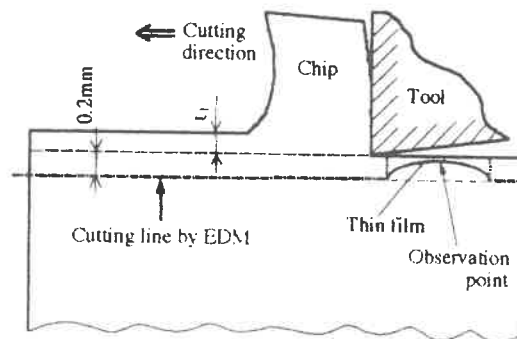


Fig. 1 Observation point of specimen in workmaterial

Table 3 Properties of machined surfaces

		Ultra Precision cutting			Precision cutting	
		$t_1 = 2\mu\text{m}$	$t_1 = 4\mu\text{m}$	$t_1 = 8\mu\text{m}$	$t_1 = 1\mu\text{m}$	$t_1 = 3\mu\text{m}$
Tangential force N		0.31	0.22	0.54	4	8.3
Normal force N		0.24	0.08	0.48	1.8	2.5
Feed force N		0.03	0.04	0.02	—	—
Microvicker's hardness Hv		64.9	—	—	27.6	33.8
Surface roughness	Ra nm	2.4	2.4	2.4	150	230
	Rmax nm	10.0	10.0	14.9	750	2920

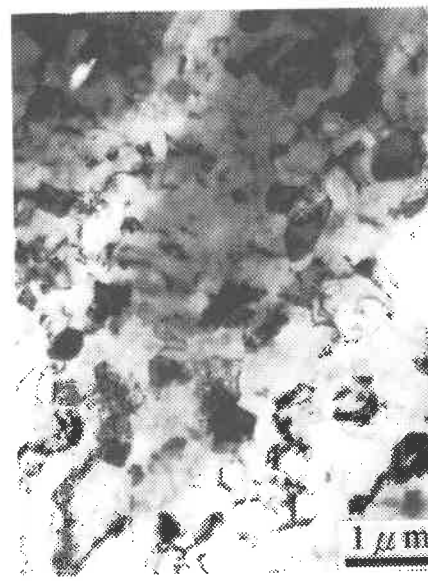
The ultra precision machinings positively provided much more sophisticated values in surface roughness. It is not reasonable to compare the cutting forces between the precision and ultra precision cuttings, because the ultra precision cuttings were not orthogonal cutting. In the same manner as the cutting forces, the hardness in the ultra precision machining was higher than in the precision machining, for the workmaterials were an Al-Mg alloy metal.

3.2 TEM observations

Fig.2 shows a TEM photo of the surface microstructure machined with 0° rake angle cemented carbide tools as a typical conventional machined surface microstructure to compare with the precision and ultra precision machined surfaces. Depth of cut t_1 was $10\mu\text{m}$. A TEM image shows a number of small polygons with high contrast. These polygons were recognized to be subgrains because of their shape, size and contrast. As the plastic deformation increases in the primary deformation zone of the workmaterial, dislocations multiplied remarkably by a frequent cross slip that easily occurs in deformation of high-stacking fault energy-fcc metals like Al, Cu, etc. Subsequently, the dislocations tangled and formed cell structure. Burnishing at the tool clearance face allowed the cell structure to dynamic-recover. This resulted in a variety of subgrains. Severe directional plastic deformation in the primary shear zone and the burnishing caused the elongated structure. High contrast in each subgrain resulted from their disorientation. This depended on whether the electron beam, penetrating into the thin film, meets the Bragg condition or not.

There are a few low contrast bands along the cutting direction in the photo. Difference in the film thickness inevitably produced the contrasts in the image. These appear to be the tool marks caused by dropping away of the tungsten carbide particles of the tool edge.

Electron beam analysis provided not a halo but a diffraction ring. A beilby layer, therefore, was not formed on the machined surfaces with these cutting conditions. Selected area diffraction analysis demonstrated almost all diffraction patterns represented a (110) spot, hence a (110) texture appears to be formed



$$\alpha = 0^\circ, t_1 = 10 \mu\text{m}$$

Fig.2 Typical conventional machined surface microstructure by cemented carbide tool



Fig. 3 Precision machined surface microstructure as presented in rolled ductile metal plates.³⁾

Fig. 3 shows a TEM photo of the orthogonal precision machined surface microstructure with a diamond tool having 0° rake angle, and depth of cut $t_1 = 1\mu\text{m}$.

The microstructure indicated that a number of subgrains produced were elongated along the cutting direction as same as the cemented carbide tool machined surfaces. No significant differences were recognized between the precision and conventional machined surfaces despite of different tool material and cutting conditions. However, the precision machinings offered



Fig.4 Ultra precision machined surface microstructure

no contrast bands caused by the tool edge irregularities.

Electron beam analysis also suggested a (110) texture and no beilby layer on the machined surface.

Fig.4 shows a TEM photo of the ultra precision machined surface microstructure with depth of cut $t_1=2\mu\text{m}$. The TEM image was slightly different from the others. That is to say, a variety of polygonal subgrains with high contrast could be found on all over the images in the precision and conventional machined surfaces, while, most of the microstructure of the ultra precision machined surface developed into not subgrain but cell structure. However, some subgrains could be obviously observed in the image partially. Therefore, the microstructure in the ultra precision machined surface was a kind of transient stage from cell to subgrain structure. The ultra precision machinings produced less damaged surface layers than the conventional and precision machinings.

3.3 Frequency rate of subgrain diameter

Fig.5 shows frequency rates of subgrain diameter in the machined surface including cemented carbide tool machined surfaces, however, except the ultra precision machined surfaces. Even though the rake angle and the depth of cut varies, the most frequent rate is undoubtedly $0.1\mu\text{m}$, accounting for more than 43% at the total. Hence, about half of the subgrain size is $0.1\mu\text{m}$. These five curves have almost the same configuration, and therefore, the subgrain size did not appear to depend on the depth of cut or tool

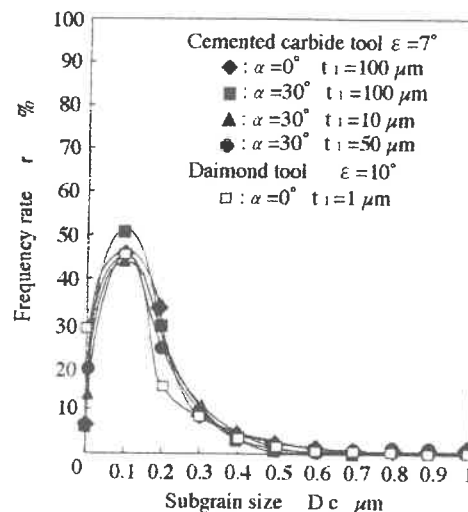


Fig. 5 Frequency rate of subgrain size

material.

4. Conclusions

- (1) Tangled dislocation networks were dynamic-recovered to produce a number of subgrains elongated along the cutting direction in all but the ultra precision cuttings.
- (2) Most of the microstructure in the ultra precision machined surfaces developed into not subgrain but cell structure, therefore, the ultra precision machinings produced less damaged surface layers than conventional and precision machinings.
- (3) Depth of cut, rake angle and tool material effects on the diameter of the subgrain could not be found, hence the subgrain size was $0.1\mu\text{m}$ for all but the ultra precision machining.

References

- 1) Kaneeda, T., Kawabe, H., Ikawa, N., & Tsuwa, H.: Microscopical Separation Process at a Tool Tip in Metal Cutting, Bull. of the Japan Soc. Prec. Eng., 17, 1 (1983)25.
- 2) Kaneeda, T.: Spreading Effect in Ductile Metal Cutting, Advancement of Intelligent Production, Elsevier Science B.V. (1994)385.
- 3) Horiuchi, S., Okubo, T., & Gokyu, I.: A Note on Rolling and Recrystallization Textures of Face-Centered-Cubic Metals, J. of the Japan Soc. of Metals (in Japanese), 29, 6 (1965)589.
- 4) Shouwaki, K.: Micromechanism of Ductile Fracture for fcc metals, J. of Japan Soc. for Technology of Plasticity (in Japanese), 13, 135 (1972) 237.

Statistical methods for process identification by surface classification using vision system

M.B.Kiran

B.Ramamoorthy

V.Radhakrishnan

Manufacturing Engineering Section
Department of Mechanical Engineering
Indian Institute of Technology
MADRAS, INDIA 600 036

1 Introduction :

Understanding of the significance of surface finish in the performance of components and the ability of the manufacturing processes to provide different finish have enabled the designer to specify the required finish on the part which could readily be achieved in manufacture. The most popular method of surface finish evaluation is by stylus instruments which is only post process inspection [1,2]. Increasing thrust towards automation in manufacture has over the years focused the need to automate the measurements done on the product both for production as well as for process control. Surface finish is one such parameter which is to be monitored for higher levels of automation. To overcome the drawbacks of the stylus instrument, electro-optical methods were developed [3,4]. Though the speed of inspection was good, they don't provide sufficient information for full characterisation of the surface. In this context, it was felt that machine vision could be tried for the evaluation of surfaces. Though, the approach is to capture the image of the surface by a relatively low cost vision system and to analyse it, there is a need to adopt a specific procedure for getting the surface image using different surface illuminations depending on the nature of finish. In other words, the illumination is to be selected based on whether the surface is coarse, medium or fine. This in turn requires the identification of the process. Hence, the vision approach explained in this paper finds out the process of manufacture of the surface in a broad fashion and then proceeds to analyse the surfaces based on the specific approach best suited for that process.

2. Texture analysis and Process identification

The texture identification consists of the following steps :

- a. Extraction of texture features.
 - b. Selection of texture features.
 - c. Method of texture discrimination
 - d. Classification and results.
- a. Texture feature extraction

To discriminate images with different textural characteristics it is essential to extract texture features which most completely embody the information about the spatial distribution of intensity variations in each image. These features can be extracted either directly from the image gray level values or from the frequency domain.

The direct image can provide a set of texture features through co-occurrence matrix S . The element $S(i,j,v)$ is the probability of going from gray level i to gray level j given the displacement vector d . We can extract texture features using a set of features which are sensitive to edge, contrast or other image properties.

From the co-occurrence matrix S [5], one can calculate the following texture features with the six displacement vectors ; $v = (0,1);(1,0);(1,1);(0,2);(2,0);(2,2)$;

Homogeneity feature., Contrast, The correlation feature of the image texture, The sum of squares of image textures, The inverse difference moment, Sum average, The sum variance, Sum Entropy, The entropy of the image texture, Difference variance of the image texture, Difference entropy of the image texture, The measure of correlation of the image texture.

b. Selection of texture features

To determine an optimal subset of texture measures without exhaustively evaluating all possible combinations, a method of successive selection and deletion based on Wilk's criterion was used.

Let X_{igk} be the i th feature value for the k th sample of class g , where $i=1,2,\dots,m$ and m is the number of extracted features; $g=1,2,\dots,G$ (G classes) and $K=1,2,\dots,n_g$ (number of samples in class g). $N = \sum n_g$ is the total number of samples. The Wilk's statistics, a measure of class separability that depends on within class scatter matrix W , and between class scatter matrix B , is defined as [6,7].

$$= [W_{ij}]_{m \times m}$$

where

$$W_{ij} = \sum_{g=1}^G \sum_{k=1}^{n_g} (X_{igk} - \bar{X}_{ig})(X_{jgk} - \bar{X}_{jg}); \quad B = [b_{ij}]_{m \times m}; \quad \text{where}$$

$$b_{ij} = \sum_{g=1}^G n_g (\bar{X}_{ig} - \bar{X}_i)(\bar{X}_{jg} - \bar{X}_j)$$

and

$$\bar{X}_{ig} = \left(\frac{1}{n_g} \right) \sum_{k=1}^{n_g} X_{igk}$$

$$\bar{X}_i = \frac{1}{N} \sum_{g=1}^G \sum_{k=1}^{n_g} X_{igk}$$

are the means of the class g and the total sample mean for the i th feature. The sum of within and between class scatter is the total matrix T .

$$T = W+B = [T_{ij}]_{m \times m}; \text{ where}$$

$$T_{ij} = \sum_{g=1}^G \sum_{k=1}^{n_g} (X_{igk} - \bar{X}_i)(X_{jgk} - \bar{X}_j)$$

the Wilk's statistics is the ratio of within class scatter to total scatter.

$$= \frac{|W|}{|T|}$$

We can test the classification ability of a set of m features by testing the null hypothesis of equal means.

$$H_0 = \mu_1 = \mu_2 = \dots \mu_g$$

where μ_g is the mean vector of class g .

Because of difficulty in obtaining the distribution function for U , the F statistics is taken as approximation.

$$F(G-1, N-G) = \left(\frac{1-U}{U} \right) \cdot \left(\frac{N-G}{G-1} \right)$$

A smaller U corresponds to a larger value of F , and a better feature set in terms of more discriminable classes.

The successive selection and deletion process deletes features which don't add significantly to class discriminability and adds successively those which do, until the most effective set is obtained. The scatter matrices are calculated using a recurrence formula which allows the determination of the incremental effect of individual features.

c. Method of Texture discrimination.

After selecting a feature subset according to the successive scheme, k -nearest neighbour classifier scheme for a set of features was implemented. Each feature was normalised by its mean and variance.

$$X_{ij} = X_{ij} - \frac{\bar{X}_j}{S_j}; \text{ where,}$$

$$\bar{X}_j = \left(\frac{1}{n} \right) \sum_{i=1}^n X_{ij}$$

$$S_j = \left[\left(\frac{1}{n} \right) \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \right]^{\frac{1}{2}}$$

Of a total of n samples one sample was held out to classify. Let X_t be the test sample and X_i be the design or training sample. The Euclidean distance between the test sample and design sample in the normalised feature space was used to determine the k - nearest neighbors of X_t . The classifier then assigned X_t to the class most frequently encountered in the k neighbors. The entire set of training samples was classified in this way, holding out a new sample each time. In addition, a second set of independent samples was classified according to the nearest neighbors in the training set. The value of k was varied to determine its effect on classification.

Let $(u_1, u_2, \dots, u_i)$ be a sequence of units with correspondence data sequence $(d_1, d_2, \dots, d_i)$ and known category identification sequence $(c_1, c_2, \dots, c_i)$. A simple nearest neighbor decision rule treated the units independently assign a unit u of unknown identification and with pattern measurements or features d to category c_j , where d_i is the pattern closest to d by some given metric or distance function [8].

d. Classification and Results :

Specimens are made of different manufacturing processes- Milling(M), Shaping(S), Electro-Discharge-machining(E.D.M), polishing(P), sand blasting(SB), Grinding (G). Thus there are six classes with different cutting conditions (speed, feed and depth of cut). Figure 1.0 shows two different of texture images.

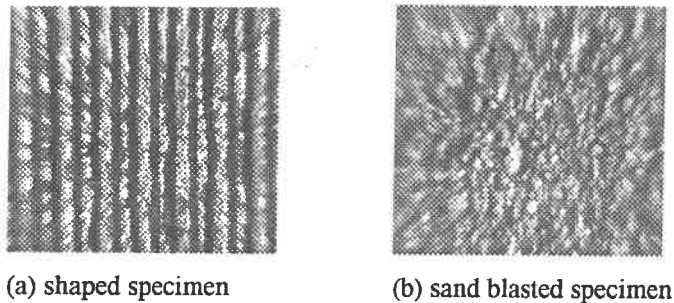


Fig. 1.0 Different texture images.

The training set for each class comprised 16 non-overlapping 32x32 sub images taken from 512x480 digitised version of the above textures.

The different features were extracted from the co-occurrence matrix.

Table 1.1 The classified Results of Training sets(1)

Real Samples of class	1	2	3	4	5	C.R.R
1	16	0	0	0	0	
2	1	15	0	0	0	
3	0	1	15	0	0	97.5
4	0	0	0	16	0	
5	0	0	0	0	16	

Table 1.2 The Within-class and Between-class Distance for Texture features.

Class	S	M	G	EDM	SB
S	0.05	0.0125	0.017	0.0026	0.0063
M	0.0125	0.0004	0.0048	0.0474	0.0188
G	0.0174	0.00486	0.0002	0.0475	0.2371
EDM	0.0026	0.0474	0.0475	0.0007	0.0026
SB	0.0063	0.01883	0.2371	0.0026	0.0006

The successive selection and deletion was applied to obtain an optimal subset of features using the image (after noise removal) texture sample training set

(16 independent samples from each of five classes) Features f1, (homogeneity feature), f2(contrast), f3(the correlation feature of the image texture), f4(the sum of the squares of the image textures), f5(inverse difference moment), f9 (the entropy of the image texture), f10(difference variance of the image texture), comprised the best feature set. Classification accuracy on the training set was above 96%. Using an independent set of 16 different sub-images for each class taken from the different regions of the original image, classification accuracy was above 94%.

2 Surface roughness assessment.

The technique is based on the analysis of scattered light from a surface. Beckmann and Spizzichino showed [9], that the microscopic waveform of the surface profile modulates the incident beams into scattered beams whose intensities and angles can be described as functions of the amplitudes and wavelengths of surface topography[10]. Thus, it is possible to obtain the surface roughness information of the surface by studying the light scattered pattern. The pattern is represented by a histogram. The histogram is a plot of frequency vs. image intensity.

Procedure : Figure 1.7 shows the experimental set-up.

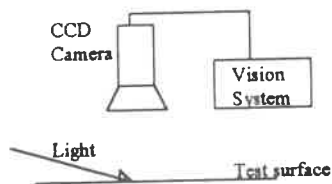


Fig. 1.7 Experimental set-up.

i). Estimation surface roughness parameter

The surface roughness parameter is denoted by R and is defined as the normalised standard deviation.

The processed image is fed to the estimation module to compute roughness parameter, R .

ii). Results and conclusions :

Good correlation was found between the vision roughness parameter and the average roughness (R_a) obtained from stylus instrument. Variability of the measured parameter was not high. Figure 1.8 shows the effects of grazing angle on vision roughness parameter. The greatest difference in roughness between two samples is obtained for small grazing angle. This leads to a conclusion that to obtain a better sensitivity a small grazing angle must be used.

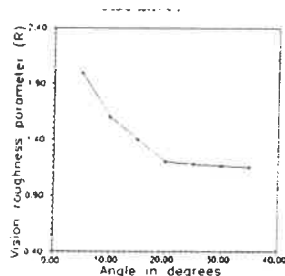


Fig. 1.8 Effect of grazing angle on vision roughness parameter.

Light scattering method holds great promise for measuring surface finish, which is non-contact, fast and convenient. Tangential lighting method is very well suited for measuring surface roughness of fine surfaces, like grinding. As the roughness increases the coefficient of variation increases indicating the spread of results and are not suitable for rough surfaces.

Its repeatability favours well with other methods.

Acknowledgment

This work is a part of the project no. MEE/92-93/084/CSIR/VRAD, funded by the Council of Scientific and Industrial Research, New Delhi, India.

REFERENCES :

1. Thomas T.R, Recent advances in the measurement and analysis of surface micro geometry., 1975., Wear, 33, PP 205-233.
2. Davis E.J., and Stout K.J., Stylus measurement techniques- a contribution to the problems to the parameter variation., 2nd Int.conf. on Metrology and properties of Engineering Surfaces., Leister, UK.,
3. Jansson D.G., and Bourke J.M., High Speed Roughness Measurement ., J. Engg. Ind., 106, 34-39, 1984.
4. Griffiths B.J., Middleton R.H., and Wilkie B.A, Light scattering for the measurement of surface finish., Int J.Prod. Research.32, 11, 2683-2694, 1974.
5. Haralick R.M., Shanmugam K., and I Dinstein, Textural features for image classification., IEEE Trans. Systems Man and Cybern., SMC-3, 6, 610-621, 1973.
6. Haralick R.M., Statistical and structural approaches to texture, Proc. IEEE 67, No. 5, 1979, 786- 804
7. Rao C.R., Linear statistical Interference and its application. Wiley , NewYork., 1973.
8. The research group of probability and statistics computation enter for the Chinese academy of sciences., The computation of probability and statistics. Scientific press Beijing., 1979.
9. Beckmann P. and Spizzichino A., The Scattering of Elecrto magnetic waves form rough surfaces.,(Oxford: Pergamon), 1963.
10. Luk F., huynh V., and North W., Measurement of Surface Roughness by a Vision System., Journal of Physics E Scientific instrumentation., 22, 977, 1990.

Highly Accurate Technique for SPM Measurement

Satoshi KIYONO, Makoto SATOH, Toshiyasu NANJYO and Wei GAO
Department of Mechatronics, Faculty of Engineering
Tohoku University, Sendai, 980 JAPAN

1. Introduction

The PZT used as a servo-actuator in the SPM probe can control sub-nanometric displacement. However, this element has considerable nonlinearity and hysteresis, which constitute different deviations from the linear relation between induced voltage and displacement resulting from the increase and decrease of induced voltage, thereby causing 3-dimensional deformation of images obtained by the SPM. This deformation is neither remarkable nor important for microscopic observation on the atomic scale.

However this distortion constitutes the greatest problem in the metrological use of SPMs for 3-dimensional measurements of ultra precision machined surfaces that have a relatively wide range of 3-D displacements.

In order to correct the distortion of SPM images, the software method which compensates measured data using the stored calibrated hysteresis curve, may be used[1]. The other method employs direct measurement of displacement of the actuator[2]. These methods are effective for correcting distortion of the scanning actuator. However, distortion of servo actuation is not easy to correct using these methods.

The other important measure for improvement of accuracy of the SPM is the development of a calibration system for nanometric displacement of the probes. In the present study, a new method of self-calibration of the servo actuator of the SPM was developed. The method can accurately determine hysteresis curves of the PZT without the aid of any other reference.

2. Self-calibration of hysteresis curve shape

2.1 Principle of Self-calibration of the scanning probe

Self-calibration is based on the same principle which has been successfully employed for the calibration of capacitance-type displacement sensors and interferometric displacement sensors [3]. Since the purpose of calibration is to compensate measurement data by using calibration data that have been obtained beforehand, it can be assumed that when the pattern of input voltage variation is repeatedly given, the output pattern of displacement of the SPM

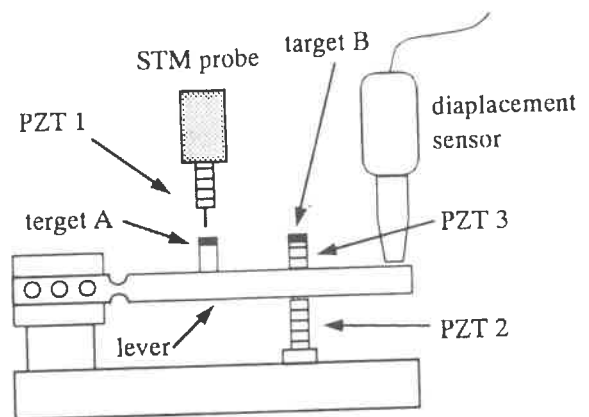


Fig. 1 Lever system for self-calibration

probe (PZT actuator) will repeatedly appear. Assuming this repeatability, the basic calibration system of the SPM probe is constructed as in Fig.1. Three kinds of PZTs are included in Fig. 1: a probe servo-actuator PZT1, a lever driving PZT2, and a zero-shifting PZT3. In the conventional self-calibration method for displacement sensors, one of two displacement sensors is placed at the reference position and the other is placed at the calibrated position. Here, however, one sensor is alternately placed at both positions and the input-lever is repeatedly driven, two times per cycle. The displacement sensor is used to check repeatability of PZT2 and mean sensitivity of PZT1, neither of which can be determined independently.

Figure 2 shows a schematic diagram of the self-calibration process for a SPM probe. The input displacement ratio of the reference point to the calibration point is set at $n=2$.

(1) First, as seen in Fig.2 (a), the input displacement of the reference point B is measured by the SPM probe, whereby the given mean sensitivity is used to estimate the displacement from the driving voltage of the actuator PZT1. When the displacement reaches the upper limit displacement of the probe, the zero-shifting PZT3 is extended and the probe again assumes the lower limit value. The estimated displacement includes the unknown linearity error of

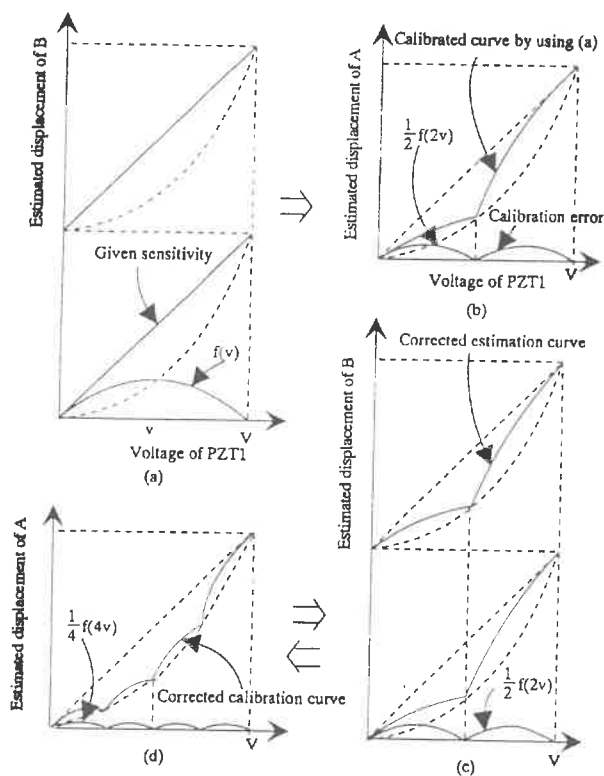


Fig. 2 Principle of self-calibration

the probe.

(2) In the next step, as seen in Fig.2 (b), the probe is set at the calibrated position. The lever-driving actuator PZT2 is extended once more and displacement of the calibrated position is measured. Dividing the measured displacement of the reference point by $n(=2)$ and using the result as the reference input at the calibrated position yields the calibration curve of input displacement induced by output voltage, which includes one-half of the original unknown linearity error.

(3) In the next step, the calibration result obtained in step (2) is applied to the data of step (1), and the estimated displacement at the reference point is compensated. {Fig. 2 (c)} Subsequently, estimated displacement at the calibration point in step (2) is compensated. {Fig. 2 (d)}

(4) Step (3) is repeated until the correction of the calibration curve becomes less than the given allowance.

2.2 Convergence of self-calibration

As mentioned previously, in the present method mean sensitivity of the calibrated actuator is given externally. The method is used

to determine the deviation of the calibration curve from this given linear line. Let this unknown deviation be $f(v)$. Displacement at calibration point A, estimated from the reading at reference point B, is expressed as follows:

$$X_{AB1} = \{X_B + f(v_{m1})\} / n = X_A + f(v_{m1}) / n. \quad (1)$$

$$V_{m1} = \text{mod}[nV, V_p], \quad 0 < V < V_p$$

where V_p is the maximum voltage applied to PZT1 and v_{m1} is the fraction of nV/V_p .

Directly measuring the displacement at point A, and comparing it with the estimated displacement in Eq. (1) yields the calibration curve shown in Fig.2 (b). Linearity error in terms of the estimated value X_{AB1} is given as follows:

$$f_1(v) = X_{AD} - X_{AB1} = f(v) - f(v_{m1}) / n. \quad (2)$$

where X_{AD} is the reading at point A, and $-f(v_{m1})/n$ is the unknown error.

By using $f_1(v)$, the displacement X_B and subsequently the displacement X_{AB1} can be corrected. After repeating this correction for k iterations, the k -th evaluation of linearity error is expressed as follows:

$$f_k(v) = X_{AD} - X_{ABk} = f(v) - f(v_{mk}) / nk \quad (3)$$

$$v_{mk} = \text{mod}[nkV, V_p]$$

Thus, the unknown term $f(v_{mk})/nk$ approaches zero with increasing k .

3. Compensation method for linearity errors

3.1 Characteristics of hysteresis loop

Figure 3 shows the results of self-calibration with a driving voltage of 40-60 volts. The figure shows both the hysteresis loop and the linearity error. Deviation from linearity is seen to approximate a parabolic curve. The maximum linearity error is 1% and 3% of maximum displacement in extension and retraction, respectively.

Figure 4 shows five calibration results with a constant voltage swing of amplitude of 20 v for different values of starting voltage of the swing, which is called the offset voltage. This figure also shows mean sensitivity in relation to offset voltage. The size of the loops and the mean sensitivity are seen to vary with offset voltage. For each offset voltage the loop is produced with a repeatability error of less than 1% at the maximum linearity error of the loop.

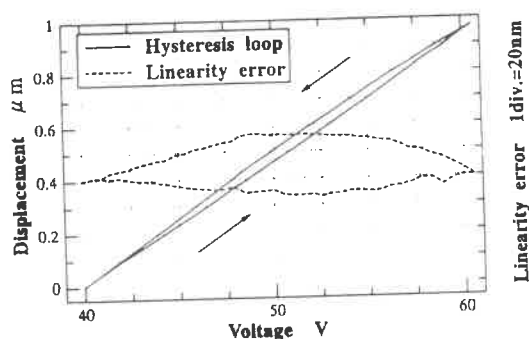


Fig. 3 Example calibration results
(Voltage amplitude: 20V)

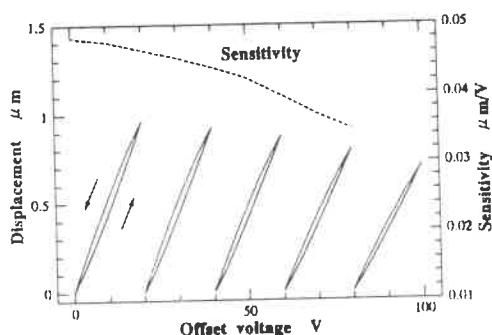


Fig. 4 Variation of hysteresis loops

3.2 Software correction.

This method uses previously obtained calibration results to correct the image after measurement. This method is easily applied to the correction of distortion in the scanning directions in which the direction of the driving voltage is clearly known.

Movement of the SPM probe in the z-direction cannot be predicted before measurement. However, in the case where the swing width of voltage is small, correction is made for only sensitivity variation since the linearity error is very small.

3.3 External monitor

Figure 5 shows the basic structure of another compensation method referred to as the 'external monitor' method. This method is based on the fact that PZT actuators from the same class show similar hysteresis patterns.[4] As shown in Fig. 5, the same voltage signal used to control the servo actuator is input to the driving circuit of two other PZT actuators. Total displacement of the two actuators is multiplied by about 5 by means of a mechanical lever, and amplified displacement is measured by a sensor. Real displacement of the SPM probe is expected to be linearly amplified at the monitoring point of the lever and detected by the displacement

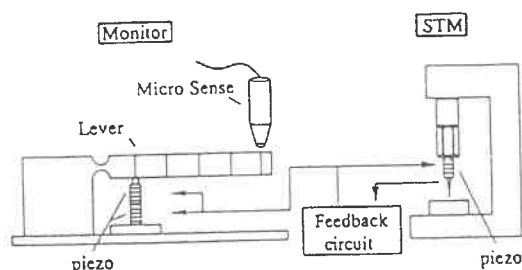


Fig. 5 Principle of monitoring method

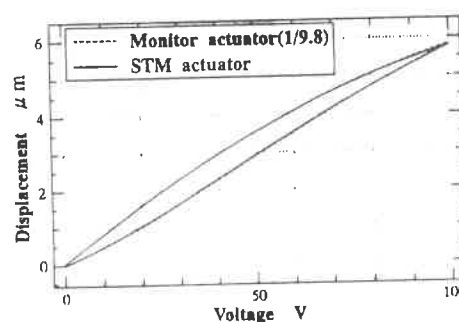


Fig. 6 Hysteresis of STM probe and monitor point sensor.

Since the resolution of displacement sensor is rough, about 10 times that of the developed STM probe, the accuracy of this method depends on the similarity of the hysteresis patterns of the servo PZT and the monitor PZT.

Although the sensor also has some linearity error, which is about 100 nm in the range of 20 μm , this error can be compensated for using the self-calibration data [3].

Figure 6 shows calibration curves of the STM actuator and the monitoring point. A magnification coefficient of 9.8 is used so that the monitoring curve can be used to estimate the STM displacement as accurately as possible. The difference between the two curves is less than 1% of the maximum displacement of the curve.

4. Results of compensation

As shown in Fig. 7, the results become almost the same as those obtained using the monitoring method. The maximum compensation is approximately 7% of the profile height. In the next step, the specimen is tilted to represent a situation that sometimes occurs during actual measurement. During actual measurement, the inclination of the specimen is removed during data processing after measure-

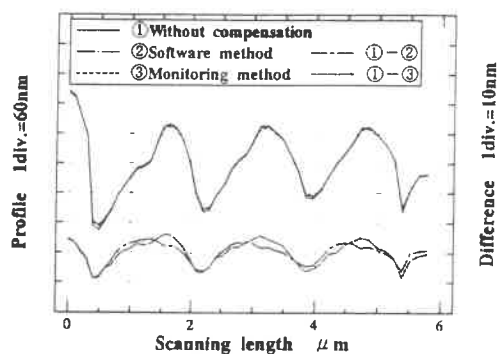


Fig. 7 Results of two compensation methods
(In software method, nonlinearity is considered)

ment. Figure 8 shows the results obtained for a tilted specimen. After the tilt is removed, not only the compensation amount but also the tendency of the compensation curve of the software becomes different from those obtained by the other method. This discrepancy is attributable to the linearity error that results from the large swing in the PZT induced by the tilt.

The displacement range of the STM image is approximately 1 μm , and the voltage swing is approximately 20V. Based on self-calibration, in this case the linearity error is known to be 1% in the direction of elongation and 3% in the direction of retraction. Inclination in Fig. 8 is in the direction of retraction.

Figure 9 shows the results of compensation of both the hysteresis loop of the large swing due to the tilt and the small swing due to each pitch of the grating. The software compensation is recognized as having approached the output of the external monitor. Thus, the two different compensation methods yield almost the same results. This indicates that both methods are reliable and offer good reproducibility.

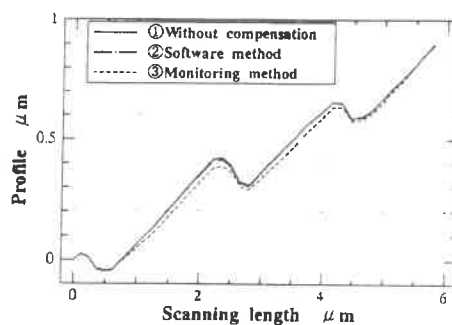
5. Conclusions

The results are summarized as follows:

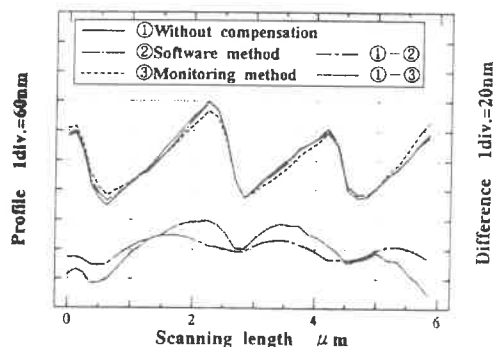
1) The proposed method can calibrate hysteresis patterns of the servo-actuator of an SPM probe to the level of accuracy of the noise level of the probe.

2) 3-D distortion of SPM images is caused not only by linearity error due to hysteresis of the servo-actuator, but also by the sensitivity variation resulting from offset voltage of the hysteresis loops.

3) Two correction methods for distortion were proposed: the software method and the external monitor method. The two methods presented similar ability for compensation amount and tendency, thus assuring reliability of the results.



(a) Without tilt compensation



(b) With tilt compensation

Fig. 8 Compensation in case of tilted sample

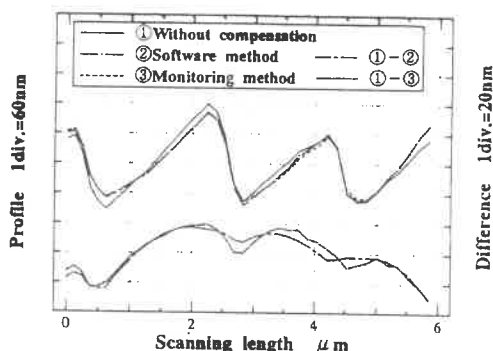


Fig. 9 Improved compensation results

4) The self-calibration method and the software correction method can be easily combined for the purpose of compensating distortion in the scanning directions. However, external monitoring method is more feasible for the correction of distortion in the z-direction.

References

- 1) R. C. Barret & C. F. Quate: Rev. Sci. Instrum. 62, (1991), 1393.
- 2) Sasaki A. *et al*: Jpn. J. of Prec. Engr. Vol. 59 (1993), 95.
- 3) Kiyono S. *et al*: Proc. of ASPE, Oct.18-23 (1992), 305.
- 4) Uchida H. *et al*: Jpn. J. of Prec. Engr. Vol. 55 (1989), 179.

A 120mm x 120mm Area Photomask Metrology Standard with Absolute Position Accuracy of 60nm

M B McCarthy¹ and A E Gee²

1 Centre for Mechanical and Optical Technology
National Physical Laboratory, Teddington, Middx., TW11 0LW, UK

2 School of Industrial and Manufacturing Science
Cranfield University, Cranfield, Bedford, MK43 0AL, UK

INTRODUCTION

The photo-lithographic process of very large scale integrated (VLSI) microelectronics fabrication is based upon the presentation, placement, photoreduction and fine movement of pattern-masks to generate sub-micrometre features on planar microcircuit 'wafer' substrates. For this, two major issues of accuracy arise: the two-dimensional positioning of features on the photomasks and the precise (stepper) motions which are essential if transfer of patterns from photomasks to wafers are to meet current overlay accuracy requirements of <100nm. (Boworth et al 1995) This has considerable implications for the metrology of artefacts and equipment and, as part of UK (NPL Teddington), FRG (PTB Braunschweig) and European Community (BCR Brussels) standards program for the development of a two dimensional high precision photomask standard, an appropriate areal metrology facility was required.

To fulfil this requirement, at NPL a number of design approaches were considered, including in-house special purpose X-Y metrology instrument design/construction and, alternatively, employing existing instrumentation. Investigations to determine an optimum solution led to a decision to modify and enhance an NPL Leitz two-dimensional optical photomask comparator and combine its measurement results with those taken using a recently upgraded NPL 400mm range single-axis rectilinear machine (the "400mm machine") using a specially-developed Gauss-Newton mathematically iterative approach.

Details are given regarding the design of the NPL 400mm machine contributing to an absolute accuracy of 1 part in 10^7 and modifications to the comparator are discussed. Graphical results from the linear and area measuring machines are given and of their mathematical combination, achieving state-of-the-art capability for the calibration of seven-inch two-dimensional photomask standards (uncertainty of 5 parts in 10^7).

NPL LEITZ PHOTOMASK (AREAL) COMPARATOR

The Leitz Model MVG7X7 Automatic Area Comparator was employed for the comparative inspection of the position of features on large scale integrated circuit photomasks. It is based on a Y-stage (upper stage) stacked on top of an X-stage, with overhead viewing objectives some 200 mm apart. It can traverse its stages in both directions while holding the edges of a pair of photomasks 50mm apart. The 200mm separation of the objectives incurs a large Abbe offset error, considerably affecting the performance of the 'Y' measurement axis. Also, small temperature changes can cause the roll, pitch and yawing motions of its stages to vary. Such deviations of the stage motion became apparent when the long term reproducibility of the machine was evaluated using the manufacturer's historical data calibration approach which employs a (13 x 13) matrix of corrections, corresponding to ~10mm increments of motion in either axis (Leitz-Anon 1988).

To replace the commercial 'historical' stage correction a 'real time' approach has been developed at NPL. For every measurement-scan performed by the comparator during any normal measurement run, the extraneous yawing motion over a 120mm range of its upper stage is measured. This 'live' calibration data is used to correct for the Abbe offset error, thus alleviating the need for regular machine calibration outside of normal measurement runs.

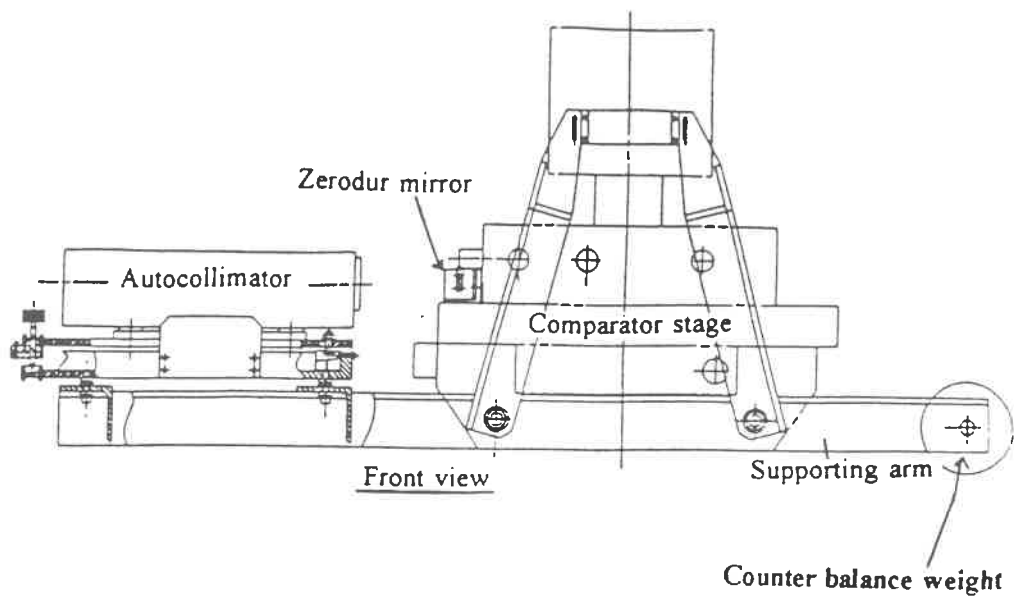


Figure 1. Commercial photomask comparator modified by NPL

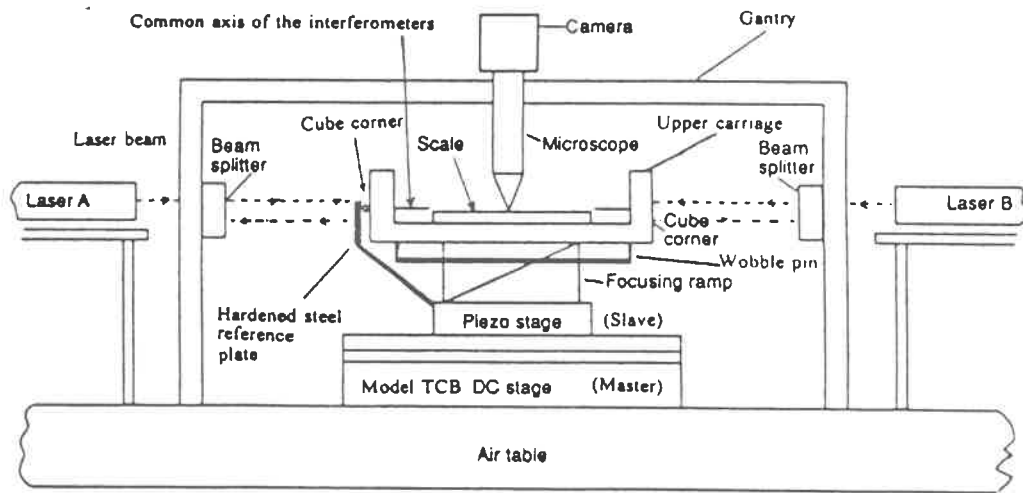


Figure 2. NPL high precision rectilinear machine ('400mm Machine')

An autocollimator is rigidly mounted to a custom supporting arm on the cast iron bed of the comparator and is used to measure the yaw of the Y-axis stage in conjunction with its return mirror attached to the upper stage. The photomasks were interchanged systematically through each of eight combinations of mutually aligned positions and the instrument brought to repeatability values of 20nm in X-axis and 60nm in the Y-axis. These repeatability values considerably improved upon the manufacturer's stated (standard) calibration specification of 120nm and 200nm respectively (Leitz-Anon 1988).

NPL HIGH PRECISION RECTILINEAR MACHINE AND AUTOMATION OF FEATURE SETTING

Principle of machine design

A fully automatic 400mm range rectilinear measuring machine has been developed for the calibration of line scales and photomask standards (McCarthy, 1995). As shown in figure 2, the machine consists of a reading head viewing a scale mounted on a master/slave stage arrangement. Scale motion is measured by a pair of Michelson type laser interferometers (McCarthy & Gee, 1992), designed taking due account of Abbe offset principles. More recently its optical reading head has been mounted using a piezoelectrically actuated flexure hinge permitting fine focus adjustment ($\pm 10\mu\text{m}$). Any extraneous lateral motion of the reading head ($< 15\text{nm}$) is monitored by a third interferometer, an NPL miniature arrangement offering nanometric resolution (Downs & Rowley 1993).

Automation of feature setting

The original method of operation used for the 400mm machine required the positioning of a feature between two cursor lines by manually controlling its stages using a computer key board and joy-stick. This process was time consuming and the repeatability reliant on human judgment and long-term diligence. To overcome these problems, the feature setting method used was automated using a video camera and image analysis. Instead of calibrating the complete viewing area of the camera, a feature is moved using the slave stage until its image fits a reference datum consisting of a selected fixed grid of pixels. This datum is defined by the user as an area bounded by a pair of horizontal and vertical cursors and, to aid feature positioning, the grid is nominally calibrated using an NPL linewidth calibration standard. The importance of this approach is that calibrations are directly related to the national standard of length (a frequency), which in practice should be realised using interferometry, and not by a mixture of precisely calibrated camera data and interferometry information.

The custom feature setting system developed in conjunction with Kinetic Imaging Ltd (Liverpool. L3 4BQ. UK.) consists of a Pulnix Model TM765 CCD camera reading head based on a 768 X 581 array of pixels at 12.5 μm pitch. Video data from the CCD camera is sent to a slave PC computer which is fitted with a Matrox IP8 frame-grabber and image processing board. This slave PC is interfaced to a master PC controlling the 400mm machine. The frame-grabber stores a TV image of a feature and pixel intensity values for the feature edges are extracted and processed using mathematical splines. For discussion on such splines see Browne & Gaydecki (1987) and Barsky (1984). In the case of the NPL system, fifty inter-pixel fits are made giving a theoretical resolution in the order of 0.02 pixel (ie 0.2 μm at the reading head). Images are magnified before reaching the camera and this increases precision.

Linear machine measurement cycle

The master PC instructs the master stage to advance under closed-loop to a position where it expects to find a feature within the field of view of the reading head. Master stage positioning is achieved to an accuracy prescribed by its radial encoder and leadscrew arrangement (typically $\pm 5\mu\text{m}$). It is then servoed, using the leadscrew, to a more accurate position ($\pm 1\mu\text{m}$) with the displacement measure provided by one of the long range interferometers. Assuming the feature falls within view, the image analyser determines its position with respect to a defined datum. The master PC then requests from the slave PC (the feature setting system) the error position of the feature. Depending on the magnitude of the error, the master PC will either move the master stage, unless the scale feature is positioned to within $\pm 5\mu\text{m}$ of the required

datum, in which case the slave stage will be actuated. The master PC repeatedly interrogates the feature setting system requesting the feature error position, each time dependent on its magnitude, instructing the master or slave stage to move. Once the scale feature has been positioned to within the desired tolerance (usually $\pm 0.02\mu\text{m}$), the master PC reads the laser interferometer counts. The process is monitored using the mean fringe-count of the interferometer to provide measurements of both feature-to-feature increments and of accumulative feature position with respect to the scale zero.

120mm x 120mm AREA STANDARD

NPL has developed a two-dimensional photomask standard, its specification being derived from a study of industry technology requirements and forecasts. NPL X-Y position standards are widely used to calibrate two co-ordinate measuring machines, photomask comparators and to aid the verification of pattern generators. The photomask standard is also used to convert a photomask comparator into an absolute measurement machine within its own capabilities.

The photomask standard consists of a quartz plate (of dimensions 177.0 mm x 177.0 mm x 6.25 mm) coated with anti-reflective chromium. The plate bears a grid design produced using an excimer laser pattern generator (ETEC-Anon, 1994) and consists of two orthogonal sets of seven parallel lines spaced at approximately 20 mm. At each of the 49 line-intersections, a chromium cross shaped feature is located, consisting of two 100 μm long lines each 5 μm wide.

CALIBRATION OF THE XY STANDARDS BY COMBINING DATA FROM COMPARATOR AND RECTILINEAR MEASURING MACHINE

A photomask comparison routine based on a Gauss-Newton iteration algorithm has been developed at NPL (Forbes 1991) combining comparative and rectilinear data. For the calibration of one photomask standard, an idle photomask is also required, although rectilinear measurements are only required from the idle photomask since precision length is transferred to the standard using the photomask comparator.

The following input data is required:

- 1) Eight sets of comparative measurement data for all the features using a photomask comparator. Two sets of measurements are taken with the photomask standards in each of the following orientations 0 , 90 , 180 & 270 with respect to the comparator co-ordinate system.
- 2) Rectilinear measurements of the features in every row and column are required on the idle photomask. Additionally, two rectilinear measurements are required diagonally across the features on the idle photomask.

The resultant data is processed using the 'photomask comparison algorithm' to calculate the absolute X-Y co-ordinates for the photomask standard under test.

RESULTS

Rectilinear measurements

To demonstrate the performance of the 400mm machine, a 200mm Zerodur scale bearing 200 lines of almost uniform linewidth was calibrated. The results were compared with those achieved by PTB and are shown in figure 3. NPL and PTB agree to an absolute accuracy of $\pm 20\text{nm}$ for linear measurements on a 200mm length optical scale which is equivalent to 1 part in 10^7 (Hantke, Häßler-Grohne, Pielles McCarthy 1994).

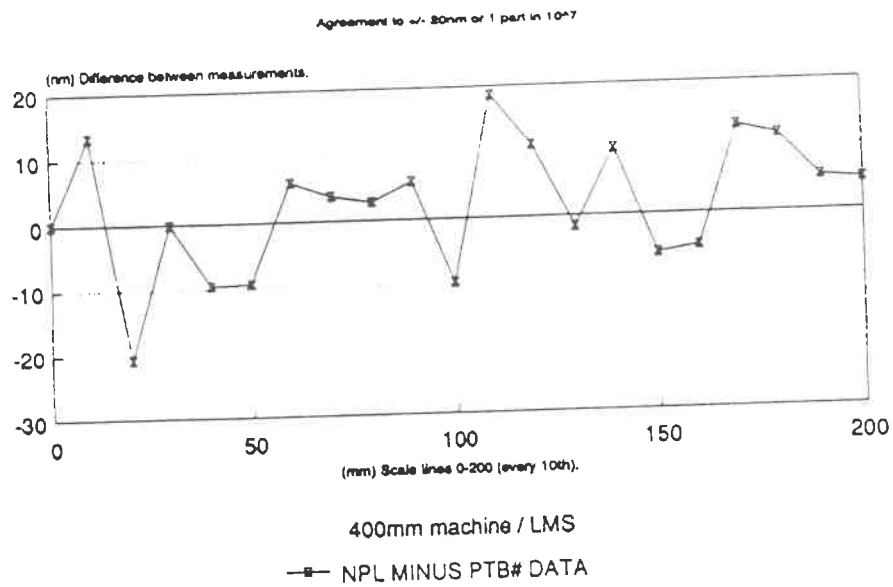


Figure 3. Length intercomparison between NPL and PTB using 200mm line scale. Measurement agreement 20nm ($1 \text{ part in } 10^7$).

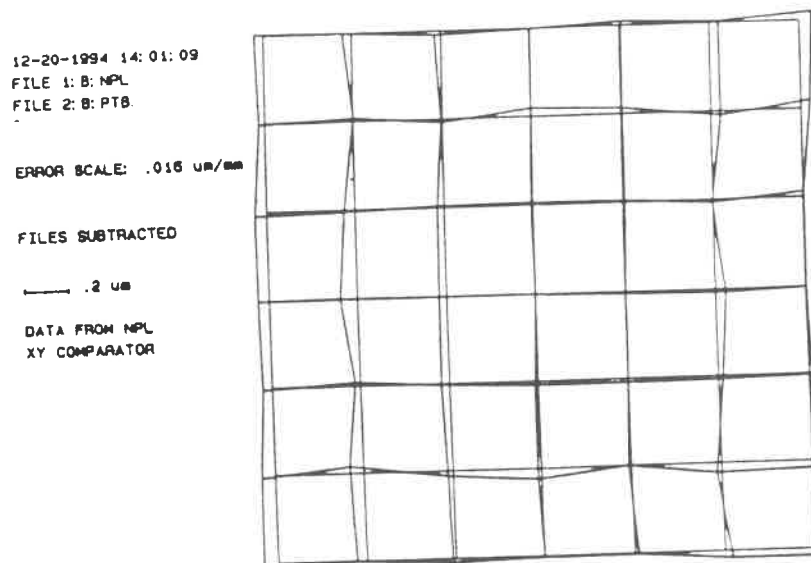


Figure 4 Two dimensional intercomparison between NPL and PTB using 7 inch photomask. Measurement agreement 60nm ($5 \text{ parts in } 10^7$).

Two-dimensional measurements

To demonstrate the performance of the NPL areal metrology facility, a seven-inch two-dimensional photomask standard was measured using the NPL 400mm rectilinear machine in conjunction with the NPL modified areal comparator and the recorded data processed using the NPL photomask comparison routine. The results from a measurement intercomparison between NPL and PTB (Hantke, McCarthy and Dearie 1995) using this standard are shown in figure 4. These agree with an uncertainty of 60nm in absolute positional accuracy over a 120mm square area (5 parts in 10^7). The uncertainty is considered to be limited by differences due to mounting distortion, a recognised problem when high positional accuracies are required (Underhill 1976).

CONCLUSION.

Performance improving modifications to the NPL 400mm rectilinear machine and a Leitz Model MVG7X7 Automatic Area Comparator have been discussed together with a Gauss-Newton mathematical approach for combining their measurement results. This has led to the development and subsequent calibration of a '120mm x 120mm Area Photomask Metrology Standard' with absolute position accuracy of 60nm.

REFERENCES

- Barsky B. A., 1984. Exponential and polynomial methods for applying tension to an interpolating spline curve. *Computer Vision, graphics & Img. Proc.*, Vol. 27, pp.1-18
- Boworth T., Fouéré J-C. and Smith N., 1995. Critical-dimensional and overlay metrology: a review. *Solid State Tech.*, Vol.38, pp.119-122.
- Browne M. A. and Gaydecki P. A., 1987. High-speed spline fitting, with application to boundary tracing in low-contrast digital images. *Comput. Biol. Med.*, Vol.17, pp.102-115
- Downs M. J. and Rowley W. R. C., 1993. A Proposed design for a polarization-insensitive optical interferometer system with subnanometric capability. *Prec. Eng.*, Vol. 15, pp 281-268.
- ETEC-Anon, 1994. Core model 2564 systems specification. K01765B
- Forbes A., 1991. Least-squares best fit geometric elements. Gauss-Newton strategy. NPL report DITC 140/89, Ed.2.
- Hantke D., Häßler-Grohne W., Pies H. & McCarthy M. B., 1994. Vergleichsmessungen an einem 200mm-Strichmaß. PTB-Mitteilungen (PTB. Braunschweig, Germany) 10/94.
- Hantke D., McCarthy M. B., and Dearie T., 1995. The development of a two dimensional high precision photomask position standard. BCR report ref: 3442/1/189/91/7/4/94
- Jones R. V., 1951. Parallel and rectilinear spring movements. *J. Sci. Instrum.*, Vol.28, pp.38-41.
- Leitz-Anon, 1988. MVG 7" x 7" automatic. Specification literature in English. Ref. No. 810-121/Engl.
- McCarthy M. B. and Gee A. E., 1992. A fully automatic high precision length measuring machine for calibrating optical line scales. *Proc. ASPE*, Vol 6, pp 309-312 (ASPE Ann Mtg Grenelefe, FLA November 1992).
- McCarthy M. B., 1995. A Rectilinear and area position calibration facility of sub-micrometre accuracy in the range 100-200mm. PhD Dissertation Cranfield University, UK.
- Underhill J. A., 1976. Overlay precision for micron lithography. *Proc.SPIE*. Vol 80, pp 85-94.

An Optical Calibrating Technique for Non-Contact Metrology

M. Taghi Mostafavi and Satish Pragalsingh
Department of Computer Science
University of North Carolina at Charlotte
Charlotte, North Carolina 28223

Abstract

This project presents results from experimental work which addresses factors related to optical calibration of non-contact metrology systems. The new technique of binary-based template matching that was used resulted in considerable reduction in computational complexity. The effects of different factors on the sharpness of the image are studied. A stable method of calibration is developed using grid plates. Known artifacts are measured to validate the calibration process. Different parameters were then measured to determine the level of accuracy that could be achieved using binary files. It has been found that, for certain applications, computational processes based on binary data can provide accurate results within the system resolution.

Introduction

In today's highly automated, quality conscious, manufacturing environment, metrology forms an integral part. While measuring soft materials, like rubber products, non-contact measurement is necessary to maintain high tolerances. This research project addresses the development of image processing methods and systems to provide appropriate data necessary to maintain high tolerances for non-contact measurement. There are different techniques that can be used to measure a given parameter. These techniques use either specific properties of the parameter measured or other physical phenomenon to detect or measure parameters [1,2,3]. For our experimental work, the technique used for acquiring data has been an optical system with simple CCD sensors. Measurements were made and different factors that affect the measurement process were studied. The main factors which affect measurement are mechanical and algorithm-specific factors. The mechanical factors are the type of lighting used, the alignment of the part, and the sharpness of the image. The algorithm-specific factors come about due to the type of computation used for measurement. For this work, a special purpose algorithm was used. This algorithm has been implemented on an Application Specific Integrated System (ASIC) chip to provide fast computational power [4], mainly for template matching processes.

Template-matching algorithms are most commonly used in pattern recognition. Yutaka [5] explains the implementation of a high speed vision system which uses a template matching algorithm to inspect part orientation. Some of the other examples are cell counting in microbiology applications; robot tracking [6], where it needs continuous template-matching to follow a moving object; label inspection; and security surveillance. In metrology, template-matching is used to locate specific regions of interest, where processing is to be done. For our experimental work, the template matching algorithm was used directly to obtain accurate measurements. One important factor that heavily impacts the accuracy of measurement is calibration. The calibration method for this project uses the same algorithm and type of lighting as is used for measuring a part. The calibration has been achieved using customized grid plates on chromoplated masks. Since the grid plates were made on masks, back lighting was used. Three different types of grid plates were made to study the factors affecting the calibration process. From the experimentation done, a set of requirements for the grid plate has been defined.

Different types of measurements were performed using the calibrated system. Initially, displacement measurements were experimented on to check if the calibrated values were "true" values. Then, a methodology to measure a part was developed. Screw threads were used for this experimental work.

Calibrations Process/Set-up

A grid plate that covers the field of view has a number of advantages, including:

- 1) The lens errors and alignment errors can be detected in a single match.
- 2) Both the x and y axes can be calibrated in a single match.

On the grid plate, the spacing between the patterns (in this case circle) is predefined. Specific grid plates have to be made for different fields of view. The field of view for this experiment was 14.5 x 11 mm. The suggested spacing accuracy is one eighth of a pixel at the expected operating resolution [7]. The grid plate was made on a chrome-plated glass mask. The accuracy of the mask was ± 1.27 microns.

Three different grid plates were made. The main aim of making the three different grid plates was to check their effect on the calibration process. The first grid plate had equal spacing. The second grid plate had crosses, instead of circles, to see if different shapes affect calibration. The width of the crosses has also made thinner (about 3 pixels wide). The final plate was made with varying spacing between the circles to see if the spacing between the circles affected calibration. The main advantages of using a masked grid plate were:

- 1) The calibration of the system was done using the same algorithm used for measuring. This removed the error which could be induced if two different algorithms were used.
- 2) The same type of lighting used in the measurement process was used for calibration, which removes the possibility of errors induced by different type of light source.

To calibrate the system, first, the grid plate was aligned such that the 2D grid plate was parallel to the x axis (or y axis) of the toolmaker's microscope. The fine angular adjustments for alignment were done by the yaw screw on the tool makers microscope (see fig. 4). An image of the grid plate was captured. The template used was a circle taken from the image itself. The binary file of the first circle on the grid plate is shown in Figure 1.

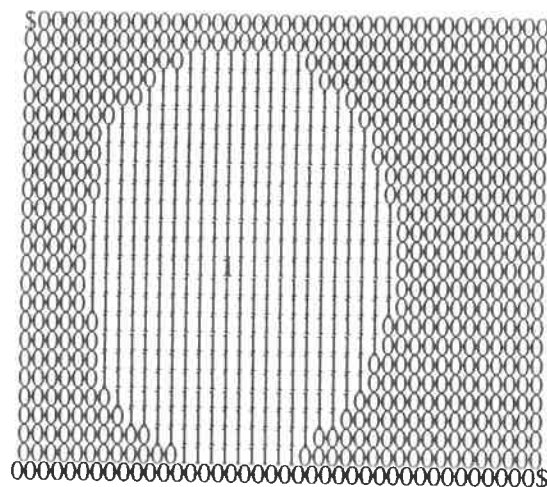


Fig. 1: Template image

Experimentation and Analysis

Using Grid Plate1, it was found that the pixel size along the x and y axes are 23.2 and 23.5 microns, respectively. The resolution in the x and y directions is 640 x 480 pixels, with a field of view of 14.5 x 11 mm. The standard deviation of the x and y axes was found to be 0.000143 and 0.000288, respectively, showing a consistent distribution of readings. For converting gray scale image to binary, different threshold values were used. The threshold values chosen did not affect calibration. While using different threshold values care must be taken that the template is also generated at the same threshold value range. The accurate calibration of the pixel size is important, as the pixel size is the unit pixel size. Any fractional error in the pixel size will have a cumulative error while measuring, as the number of pixels increases (distance). This can be seen from the units of the pixel size, which is mm/pixel. For example, in Figure 2, though the

calibrated value was 0.0232 mm, it was assumed to be 0.023 mm, inducing an error of 0.2 microns on about 0.9%. Distances were measured over the field of view of the image; a slope error was induced as seen from the measured error line. The "measured error line" is the error in measuring distances, using the calibrated value of 0.023 mm. The "slope error line" is the slope error due to the variation of 0.2 microns. The "corrected line" is the line obtained after removing the slope error.

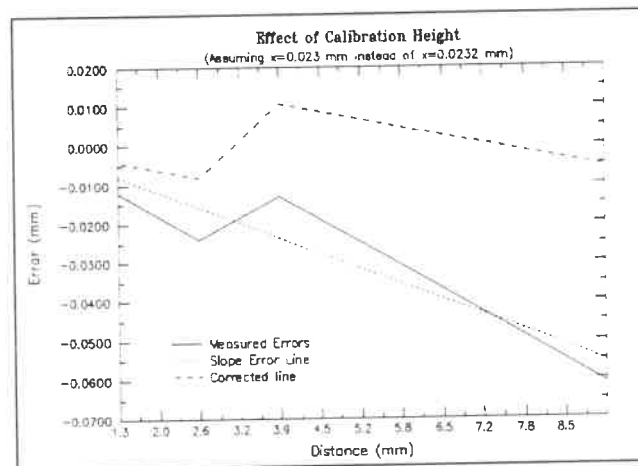


Fig. 2: Slope error of x axis

To check the calibrated values, different distances were measured using the translation stage. The distance moved versus the error of measurement along the x and y directions is shown in Figure 3. The errors are within plus or minus half a pixel. Also, there is no slope error, hence the calibrated values are "true". Grid Plate2 also produced the same result with respect to the pixel size, but since the line width of the crosses was thin, the threshold value chosen was critical. The threshold value chosen did not affect the accuracy, but would create deformed crosses, which could not be detected.

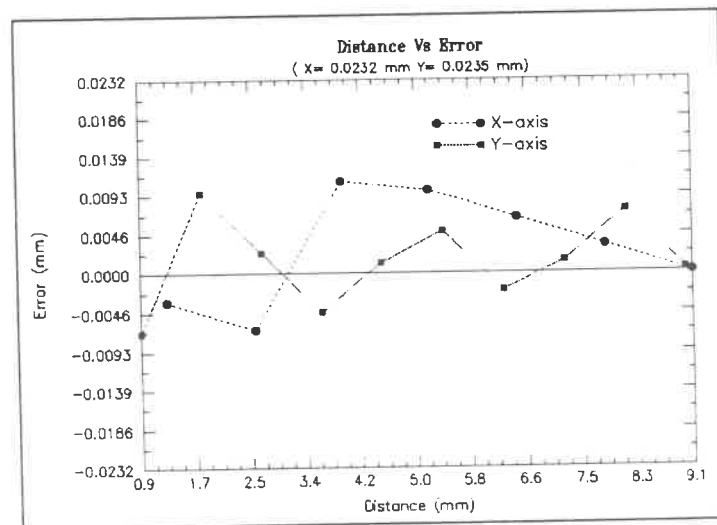


Fig. 3: Checking calibrated values

The varying spacing Grid Plate3 also gave the same pixel size for the x and y axes. Though each spacing gave a different value of pixel size when averaged over the whole plate, results similar to the previous grid plates were obtained. Each spacing was measured ten times in Grid Plate3. Compared to Grid Plate1, a given spacing was measured 100 times to produce a stable pixel size. It was concluded that 100 readings will provide stable calibration values for a given spacing.

Grid Plate Requirements

From the results obtained by using the three grid plates, the following were concluded in order to make an appropriate grid plates.

- 1) The grid plate should be large enough to cover the entire field of view easily. This requirement would check for lens errors, if any, in the periphery of the image.
- 2) The size of the patterns on grid plate should be large enough to be detected. For this, a rough idea of the pixel size can be obtained from the size of the field of view and the resolution of the vision system. For example, the circle should be greater than 10 pixels in diameter. This requirement is necessary to capture a clear image.
- 3) A pattern of 11 rows and 11 columns of circles (other shapes) is considered to be a good number of patterns. This is to have at least 100 readings to obtain a statistically-stable value.
- 4) The spacing between the circles need not be equally spaced, as long as the spacing between the circles is known accurately.
- 5) The positioning of the circles should be accurately known to at least one eighth of a pixel at the expected operating resolution.

The shape used in the grid plate did not affect calibration. Though the spacing between circles has no effect on the calibration process, having equal spacing has the advantage of detecting if the CCD array is aligned to the focal plane. Hence, the equal spacing grid plate is preferred.

Measuring Factors

To check the calibrated values, displacement measurements were performed. Afterwards a methodology for measuring screw threads was developed. A methodology to find the center of a circle was also developed. There are many factors that affect the accuracy of measurement. They are sharpness of the image, mechanical factors, and algorithm-specific factors. The first two factors affect all types of measurement. The algorithm-specific factors are the threshold value selection and the type of template used. These two factors affect the part measurement accuracy considerably. Thus, the above-mentioned factors have to be considered while developing a methodology for measurement.

To calibrate for sharpness, different readings of the same object were taken, by keeping the focal length and the intensity of light constant and by varying the aperture value. An additional parameter that was found to affect the sharpness of the image was the distance (D) between the part and the backlight. This is because, at lower apertures, more amounts of light enter the camera and blooming occurs. The effect of blooming reduces as the aperture value is increased. Hence, a sharper image is obtained. As the aperture is further reduced, a lesser amount of light enters the camera, producing a dark image. Therefore, the sharpness of the image falls again. It can also be seen that the aperture value of 11 has the best sharpness value. This value coincides with the recommended aperture value in the research work done by Randal J. Hunsicker [7] using the same setup. The optimal aperture is a property of the camera, as the effect of blooming will vary from camera to camera. The optimal aperture value must be found for the camera used. The sharpness of the image also increases as the distance D is reduced. This is because as the part gets closer to the backlight the effect of the external lighting on the part is reduced. This noise reduction produces a sharper image.

The mechanical factors deal with the physical alignments and settings of the camera and part. The three main factors are the alignment of the part, the height of calibration, and the height of measurement. See Figure 4. The alignment of the part along the direction of measurement with the CCD array is critical. For example, while measuring the pitch distance, the pitch alignment of the part with respect to the camera is important. For our experiment, the part was measured from a different plane than at which the calibration was done (see Fig. 4). Care had to be taken to ensure these two heights (H_c and H_m) are maintained within 2mm. Any variation of more than 4 mm induces an error of 0.2 microns in the calibrated value.

Measuring Displacement

Displacement measurements were performed to check if the calibrated values produced true measurements. The displacement measurement will find its implementation in the tracking of objects. To measure displacement, a translation stage was used. A single circle of the grid plate was focused and placed in the field of view. The initial position was considered the reference point and the template was created with the first position image. Then, the translation stage was moved through known distances and matched at each distance. The number of pixels between the matched points was found and the distance measured by multiplying the calibrated value. For example, say the translation stage is moved by 1.175 mm along the x axis. The calibrated value is 0.0232 mm along the x axis. The starting point is at (15,15), which is the reference point of template. The matched point after moving is at (66,15). The difference between the two values gives the number of pixels moved, which is 51. Multiplying 51 by the calibrated value the measured distance is 1.1832 mm. The error in measuring the displacement is found by subtracting the distance moved by the measured distance. For example, $1.175 - 1.1832$ gives an error of 0.007 mm.

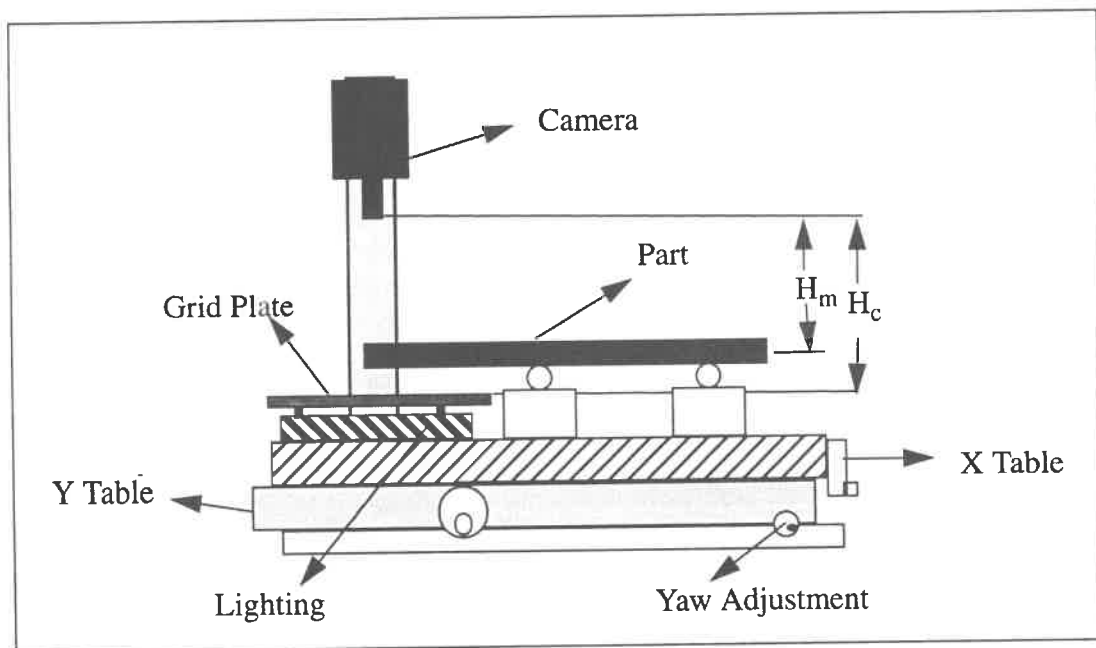


Fig.4: Mechanical factors that affect accuracy of measurement

A set of five arbitrary distances was measured four times. Each time the measurement was performed the sequence of measurement was changed. This was done to check if the same errors repeat. It was found that possible errors follow a specific pattern. For example, if a distance of 2.35 mm is measured consecutively six times, a pattern of error is followed. This is shown in figure 5 for three different readings. The pattern of error is dependent on the fractional difference between the calibrated value and the measured distance. This error can be predicted and possibly corrected. A range of ± 20 of threshold values did not affect the accuracy of measurement. This is because the measurements are done with respect to the matched points. As is characteristic of the chip, the maximum match is always at the reference point.

Consolidated Results

The use of binary files in measurement for certain applications was experimented upon. A template matching algorithm, which is conceptually similar to the generalized Hough transform [8] and uses binary images, was used for our experimental work. A stable method of calibration was developed using grid plates. Using the calibration method, the pixel size of the system was found to be 23.5 microns in the x axis

and 23.2 microns in the y axis. Measurements were made and different factors that affect the measurement process were studied. The main factors that affect measurement can be divided into two categories: the mechanical factors and the algorithm-specific factors. The mechanical factors are the type of lighting used, the alignment of the part, and the sharpness of the image. The algorithm-specific factors come about due to the type of algorithm used for measurement, including the threshold value chosen and the type of template used. Methodologies for measuring different parameters, like displacement measurement and part measurement, were developed. The accuracy of the measurements made was within the system resolution. The algorithm used does not limit the resolution of the system. For example, if sub-pixel manipulations are performed on the captured image, the same methodology can be used to produce more accurate measurements. The limiting factors are the hardware of camera, lens, and frame grabber. From the experimental work, it became clear that for certain applications, binary images can be used to produce accurate measurements.

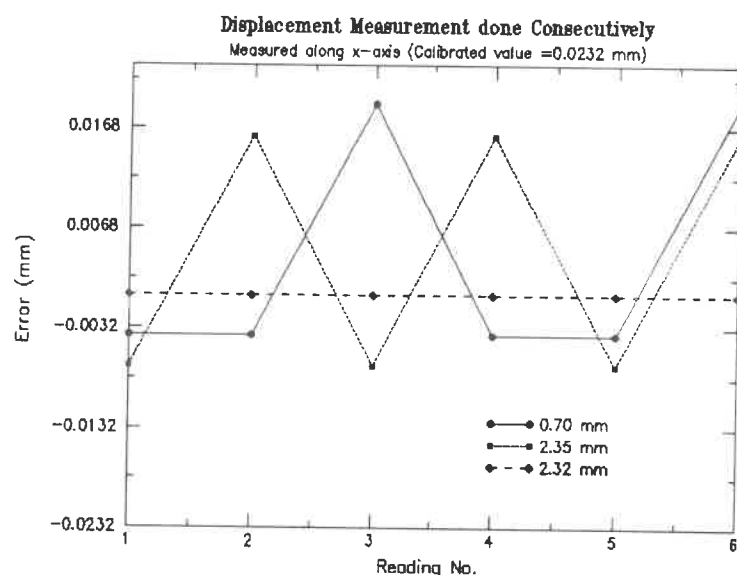


Fig. 5: Displacement measurement

Acknowledgment

We would like to thank the National Science Foundation for their funding of this research work.

References

- [1] Yiping, et al, "New Technique for Sub-Micron Line Width Measurement", SPIE AIHSVA, Vol. 849, 1987.
- [2] Stephen M., Hall, "Vision Measurement Technology Using Photogrammetry for Quality Control in Aerospace Industry", PRI of SME, Feb27-March1, 1984
- [3] Yusuke, et al "A laser Applied 3-D Profile Measuring System", H R, Vol 35 No.1, 1986.
- [4] T. Mostafavi, "A Parallel Model for Fast Computation of Pattern Matching in Integrated Engineering," PICCM in Engineering, pp 49-56, May 1993.
- [5] Yukata Ishizaka and Shoji Shimomura, "High Speed Pattern Recognition System Based on Template Matching", Fuji Electric Corporation Research and Development, Electronics Development Div., 1 Fuji-Machi, Hino-city, Tokyo, Japan.
- [6] Roger, Y., Tsai, "A Versatile Camera Calibration Technique for High Accuracy 3-D Machine Vision Metrology Using Off-the-Shelf TV Camera and Lens", IEEE Journal of Robotics and Automation, Vol. RA-3. No. 4, August 1987, pp. 323-344.
- [7] Randal, J. Hunsicker, John Patten, Alton Ledford, et al; "Automatic Vision Inspections Measurement
- [8] T. Mostafavi, "Image Processing System and Methods for Precision Metrology," P. of the 1995 NSF DMGC, UC San-Diego, pp 131, Jan, 1995.

FLATNESS AND LEVELNESS MEASUREMENT SYSTEM FOR BIG SURFACES

Rigoberto Nava S., Sergio Padilla O., Benjamín Valera O.

Centro de Instrumentos-UNAM, Apdo. Postal 70-186, Coyoacán, 04510, México, D.F.

Tel: 622 86 05, Fax: 622 86 20. Email: navar@aleph.cinstrum.unam.mx

ABSTRACT

The actual trend in warehouse construction is to occupy small areas with large volumes of storage. This causes vertical designs with high storage structures and the use of freight elevators with no much stability because the rate basis/height is little. At this situation during the construction process, warehouse floor must be evaluated to verify its levelness and flatness to make sure it complies with the required quality. In this paper we present the functional prototype of a system for flatness and levelness measurement of large areas. This system is formed by measurement instruments of length and level assembled on a light vehicle driven and guided automatically under the control of a dedicated microcontroller. The lectures at the instruments are send by radiofrequency to the main computer where they are processed with original algorithms for the analysis of levelness and flatness according to the E1155-87 ASM standard. The complete travel of the vehicle over the measurement area could be made with a high level of automation and this means a great advantage over other procedures.

INTRODUCTION

Some special floors must support traffic of vehicles for load or transport and this traffic could be aleatory or in fixed routes. We found an example for the use of this kind of floors in high vertical warehouses, these warehouses are designed to store great number of articles in a minimum area. In this cases, floors must be very flat because the area assigned to the movement of vehicles is very narrow and they could collide against the structures caused by errors in the levelness or flatness of the floor. So, it's necessary to know these characteristics to determine if the flatness deviations of the floor are into tolerances, so they can be corrected before the floor is used to prevent accidents that could cause human or material losses.

By other way, it's known that large flat surfaces (granite or iron) used for mechanical parts measuring, are calibrated in situ. One of the methods actually developed allows the calculation of the flatness by means of angle measurement, along straight lines that divide the flat surface in sections. In this case, it's necessary to spend long time to take the readings from the instruments.

Previous experience shows that acquiring and processing of data is complicated to be done manually [2]. That's why, we emphasize in the automation of the process developing a vehicle self-directed capable of perform and transmit readings by means of wireless radiofrequency.

2. MEASURING ALGORITHM

In the measuring algorithm it's described the complete measuring process divided in three basic parts: the model employed, the automatic process and the performing and processing of the readings.

2.1. Physical model of the measuring system

The model includes the position of the instruments over the surface to be measured. In this arrangement we note two sub-systems: hardware and software separated but connected through a RF communication linkage. The first sub-system is an automat vehicle automatically controlled in a straight run over the measuring area that transports instruments for length and level measurement. This controller takes the readings and sends them via RF to a central computer. The second subsystem is a RF fixed receiving basis linked to a personal computer that's waiting the asynchronous arrival of the readings for a later processing.

The algorithms in both subsystems of the model are presented in figure 1.

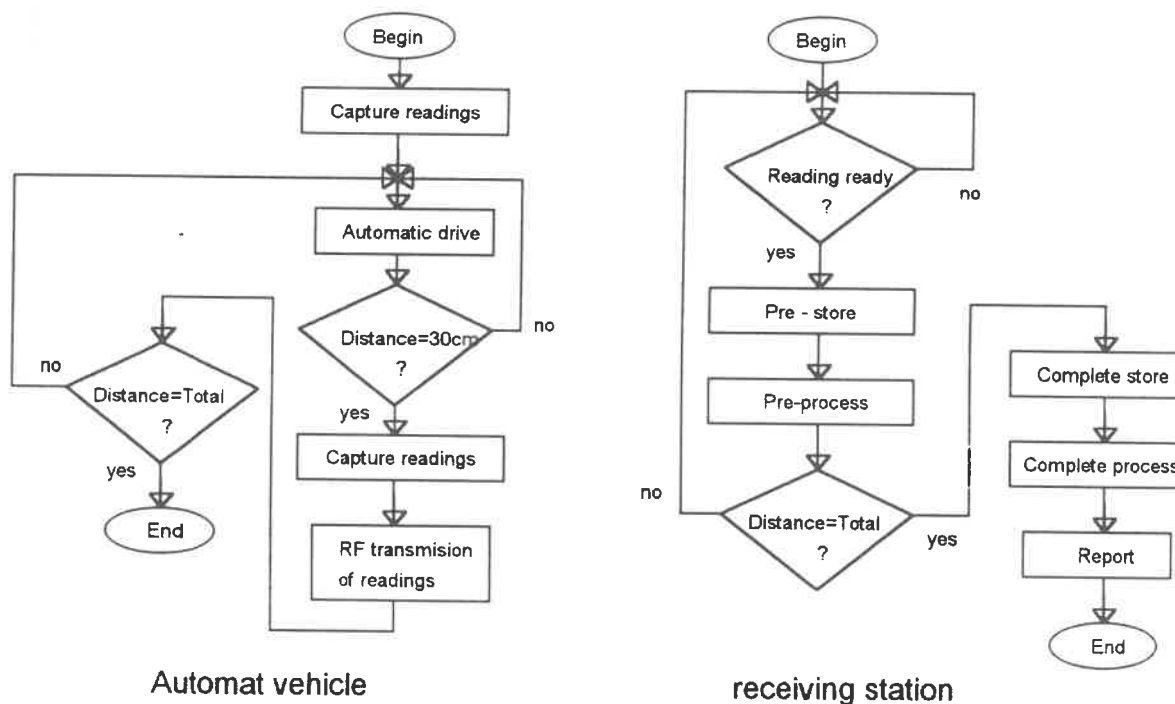


Figure 1. Algorithm for flatness and levelness measurement of extended areas.

2.2. Automatic capture of readings

The microcontroller MC68HC811E2 administrates the resources of the automat vehicle as is shown on figure 2.

The software in assembler language for the MC68HC811E2 [1] performs four main tasks simultaneously: drives automatically the vehicle on a predetermined straight path by means of

visible light contrast; captures readings from the instruments; compensates against scale overreading and then send it by radiofrequency to the host computer. The remote wireless transmission is made in basis of the asynchronous interruptions of this microcontroller.

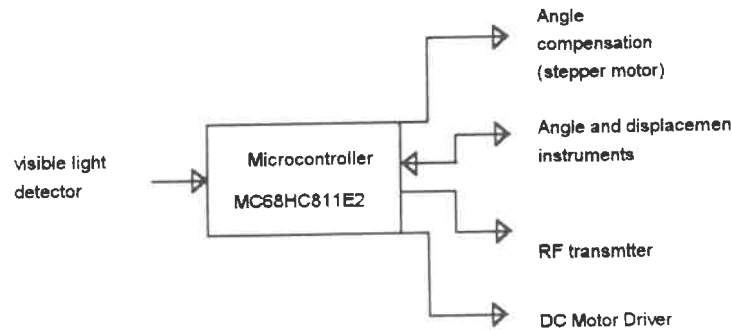


Figure 2. Subsystem of automatic drive, registering, capture and transport of readings.

2.3. Readings processing

Figure 3 shows equidistant points *A*, *B*, *C* and *D*. Points *B* and *C* are rigid supports on the surface under test, which for readings processing define a straight line segment having an angle α with respect to the horizontal line. The vertical distances *a* and *b* are deviations from points *A* and *D* in regard to the straight line *BC*.

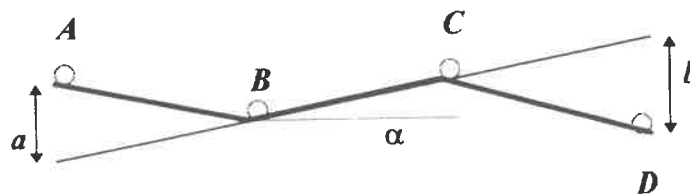


Figure 3. Data obtained from an elemental line of the surface.

With this data it is obtained the straightness of an elemental surface line. Once realized the straightness measurements from all the elemental lines previously defined from a surface, and co-relating the position between adjacent elemental lines, the topography of the total surface is obtained.

The co-relation of parallel elemental lines is made through the height change of the contact *B* point with respect to the last position in each measuring station. This change is defined by an angle β measured in a vertical plane perpendicular to line *BC* by mean of an electronic level.

The values of the angles α and β , as well as the values of straightness deviation *a* and *b* from the straight line, which passes through the supports *B* and *C*, allow to qualify the surfaces under stated standards on the respective quality for flat floor, and for flat surfaces [3,4].

3. INSTRUMENT CHARACTERISTICS

- Rigid and light structure.
- Two high precision electronic levels in perpendicular position between them.

- Two high precision electronic probes for displacing measurement.
- One display for the levels and probes with interface RS-232.
- Microcontroller dedicated to the administration of capture, transmission of readings and automatic drive.
- RF radio frequency transmitter and receiver.
- Lap-Top host computer.
- Software for analysis of flatness and levelness.

4. PARTIAL RESULTS

We have partially proved that the automatic measuring process reported in this paper exceeds widely our previous experiences [2] carried out in manual form, having enough support in the control of high precision instruments with an adequate number of significant figures. We estimate a reduction of 10 to 1 in time related with the set up and process of measurement, due principally to the data capture and control by microprocessors.

5. CONCLUSIONS

The important part of this work development is not a competition with other commercial instruments designed for the measurement of flat floors, instead of this, we try to satisfy the commitment of a price really attainable in the measurement of the parameters mentioned here in.

We put special attention in the optimization of the way it is done, considering the initial cost of the instrument, the easiness for the calibration, maintenance, transport and bringing up to date, new characteristics at the same time of its easy operation (weight and size) and integrated environment of operation.

Nevertheless, although the instrument is not thought to be handled by non qualified personnel, the easiness in the capture and processing of data is a simple work, useful and appropriate for our actual needs.

6. REFERENCES

- [1]. Motorola Inc., *HC11 Reference manual* (Motorola Inc., U.S.A., 1991), 469 p.p.
- [2]. R. Nava, B. Valera, S. Padilla, *Planimetro para medición de planitud de pisos en términos de números FF y FI, Informe técnico* (Centro de Instrumentos, UNAM), 20 p.p.
- [3] ACI Committee, *Comentary on Standard Specifications for Tolerances for Concrete Construction and Materials (ACI 117-90)*, (ACI Committee 117), 7 p.p.
- [4] ACI Committee, *Construcción de losas y pisos de concreto*, (ACI Committee 302), 5 p.p.

New NIST-Certified Length Microscale

James Potzick

National Institute of Standards and Technology
Gaithersburg, Md. 20899

A certified line scale artifact is under development at the National Institute for Standards and Technology (NIST) as a Standard Reference Material. In the form of a microscope slide, it contains a one-dimensional sequence of lines with spacings from 1 μm to 1 cm with expected placement uncertainty less than 100 nm.

1. Introduction Accurate dimensional measurements on machined parts are needed for comparing with specifications or operational models. Calibration and design verification of

microactuators and microsensors require accurate measurements of the displacement of structural components of micrometer (μm) dimensions.¹ For these purposes and others the

metrology instrument or machine tool of choice must have a calibrated length scale traceable to the meter. The only way the accuracy of a measurement can be assessed is through its traceability to a recognized standard. While the National Institute of Standards and Technology (NIST) has long provided a calibration service for customers' length scales,² there is a need for a small length scale artifact with traceable calibration. One optimized for use in all kinds of microscopes, especially metrology microscopes, would be particularly useful. When a measuring instrument is calibrated with this standard, traceable measurements are possible and certain legal requirements can be met. When a machine tool is so calibrated the part's finished dimensions will more closely conform to the specifications.

This artifact, a calibrated stage micrometer, is under development at NIST and is called SRM (Standard Reference Material) 2800, *Microscope Magnification Standard*. This standard is intended to be simple and relatively inexpensive. Much work remains to be done in completing the calibration system and documenting the calibration uncertainty. It is expected to be available in 1997.

2. Physical description This standard consists of symmetrical nested linear pitch patterns in decade ranges, with calibrated pitches (line-to-line spacing) ranging from 1 μm to 10 mm. An array of etched antireflecting chrome parallel lines is printed in one direction in the center of a quartz mi-

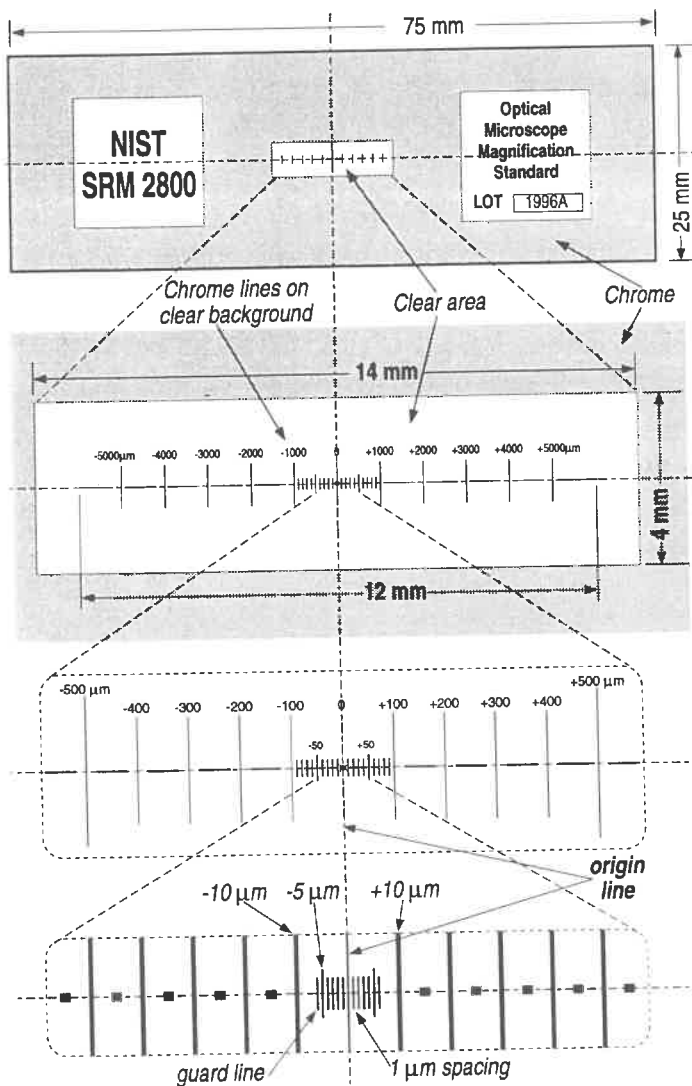


FIGURE 1 NIST SRM 2800, *Microscope Magnification Standard*. Overall view at top, successively exploded views below.

crosscope slide. The chrome lines are on a clear background to facilitate use in optical microscopes using either transmitted or reflected illumination. Figure 1 shows the overall configuration as well as details of the smaller pitch patterns. The horizontal position of the center of each line relative to the origin in the center of the pattern will be certified. The origin is at the center so the smaller lines remain in the field of view as magnification is increased.

The artifact is fabricated commercially using photomask techniques to ensure high quality. The pattern is directly written on a standard photomask blank using an electron beam or optical pattern generator, then developed and etched. The patterned photomask is then cut into 25 mm $\times$ 75 mm $\times$ 2.3 mm thick (1 in. by 3 in.) pieces and the edges finished.

3. Using this standard This standard can be used to measure the magnification of a microscope, or to calibrate the metrology scale of a microscope or machine tool. It is suitable for optical microscopes using reflected or transmitted illumination, scanning probe microscopes (SPMs), and scanning electron microscopes (SEMs). For the latter it must be coated with a thin conducting film to prevent charging by the electron beam. This pitch standard can be moved about in the field of view to measure field distortions and the linearity of metrology scales in any direction. It can be used to measure the effective pixel-to-pixel spacing throughout the field of view in microscope video systems.

The materials used are stable and robust. The calibration should last indefinitely if the standard is not damaged.

Pitch and linewidth Only the line positions, not the linewidths, will be certified. Figure 2 illustrates the difference. While this standard facilitates accurate calibration of magnification, scale, and tool displacement, care must be taken when measuring the size (left edge to right edge, length, or linewidth) of an object. The appropriate definition for "edge" becomes an important issue, and proximity effects and edge effects are important when the required measurement uncertainty is less than the wavelength of the light used in optical measurements, or less than the electron-specimen interaction volume in SEM measurements.

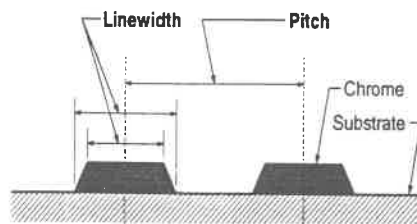


FIGURE 2. The distinction between pitch and linewidth (cross section view)

Linewidth is a much more difficult measurement than pitch and, with one exception, NIST has no micrometer-size linewidth standards because of the complexity of the mathematical modeling required for each different user artifact to be measured.³⁻⁶ That exception is for photomasks viewed in transmitted light,⁷ whose lines are all physically similar enough that one relatively tractable model can serve for the whole class.⁸

The positions of the line centers, rather than left or right edges, are certified because center-to-center pitch measurements have the advantages of symmetry, not the least being avoidance of possible confusion between right and left edges in microscopes which may or may not invert or reverse the image.

4. Calibration These standards will be fabricated commercially to NIST specifications and calibrated at NIST.

Sampling Strategy Since the available electron beam photomask writing tools are very accurate in feature placement, it is anticipated that all of the lines will be within 0.1 μ m of their nominal positions. Samples will be taken from a production lot based on a sampling strategy determined by the NIST Statistical Engineering Division, and the position of each line on each sample will be measured. Based on these measurements, the sample size will be adjusted if necessary until the probability that all of the lines in the entire lot are within 0.1 μ m of their nominal positions is 95%. The unmeasured standards will then have a certified expanded uncertainty of $\pm 0.1 \mu$ m from nominal at the 95% confidence level. The $\pm 0.1 \mu$ m tolerance may be reduced if this can be supported by the measurement data. Those samples which have been individually measured may be sold with lower uncertainties at a higher price.

Calibration instrument The sampled artifacts will be measured on the NIST Ultraviolet (UV) Scanning Microscope (Fig 3), which is also used for photomask linewidth calibrations.⁹ A precision measuring microscope has more stringent structural requirements than most other microscopes, and this instrument is based on the kinematic Stewart Platform design. The specimen stage and the optical system are mechanically connected by six lightweight struts so

that the number of constraints equals the number of degrees of freedom. The goal is to reduce internal stress in the structure and to raise the resonant frequencies of all important structural vibration modes. The struts are made of hollow aluminum tubing stock, and are connected to each other and to upper and lower mounting plates by adjustable node joints and friction clamps (patent pending¹⁰) which allow for dimensional adjustments while constraining the centerlines of the struts to intersect at the nodes. This node joint design greatly facilitates construction of properly designed Stewart platforms. The specimen stage is mounted on the lower mounting plate on an optical table, and the magnifying optics mounted to the upper plate. This results in a very rigid structure with a large working volume. There is no significant vibration damping in this structure, unlike a traditional cast iron design; the intent is to reduce relative motion between the specimen and the objective lens. The upper mounting plate initially exhibited some plate-mode vibration, so damping material was later added here. This microscope is mounted on a well damped compliant optical table as a first line of defense against vibration.

The specimen stage is constructed of stock lead-screw crossed roller bearing linear stages driven by dc motors with linear position feedback. Stage position is computer controlled in *x*-, *y*-, and *z*- (focusing) axes for positioning. The speci-

men is measured in visible or ultraviolet transmitted light by scanning the stage along the *x*-axis at constant velocity. Meanwhile, the intensity of the magnified image is measured through a pinhole sampling aperture fixed on the optical axis in the image plane, and the position of the scanning stage is measured with a laser interferometer. These measurements are rapidly repeated during the scan, building up an array of paired *intensity* and *position* data points. Scanning the specimen stage is preferable to scanning the aperture (or using a CCD pixel scan) as it provides a more direct link to the SI unit of length. Measurement accuracy is more important here than measurement speed.

Some of the computer code segments used in photomask linewidth calibrations are also used in this pitch calibration. Intensity and position data pairs collected during the scan are saved in an array in the computer to form one raster scan line. Since the maximum scan range of this microscope is greater than 2 cm, the entire line array will comprise one scan line (containing 118 edges) to minimize the time interval between successive edge measurements. A set of raster scan lines forms the data "image" of the feature. The data are digitally filtered to remove high spatial frequency noise beyond the cutoff resolution of the optical system, and fit to polynomials in the regions about the edge threshold to further reduce the effects of noise (especially vibration). The positions of the edges on each raster line are

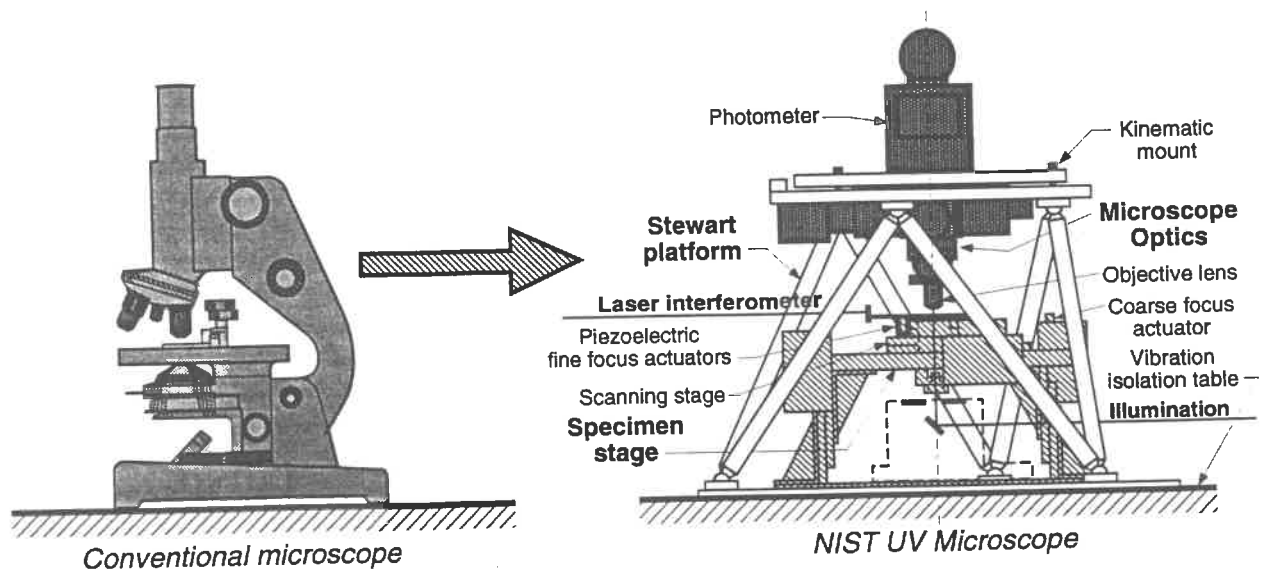


FIGURE 3. A conventional microscope (left) and the NIST Ultraviolet Scanning Microscope used to calibrate SRM 2800 pitch standards and photomask linewidth standards (right).

found using the same intensity criterion as used for linewidth measurements,¹¹ so the inferred center of the line is its physical center. The edge data for each line are tested against such quality criteria as taper or evident edge irregularities, and any alignment cosine error is removed. Right- and left-edge data are stored in a calibration file, along with measurement parameters. Then the average line center positions are calculated from the edge data. The scan data are presented graphically on screen and updated as the calculations are made.

Measurement Uncertainty The total error of a measurement¹² is the difference between the measurement result and the true value (referred to the International Standard Meter for dimensional measurements); it is the sum of systematic and random errors. The systematic error is the mean of an infinite number of measurements minus the true value; *i.e.*, the error after measurement-to-measurement variability, or measurement noise, has been removed. It is unknown and must be evaluated using all available sources of information. The random error is the result of a single measurement minus the mean of an infinite number of repeated measurements; *i.e.*, the part of the error due only to measurement-to-measurement variability.

The final measurement error is unknown because the true value is unknown (else why measure?); otherwise it could be removed from the measurement data to eliminate it. The measurement uncertainty derives from the probability distribution of the errors. The standard uncertainty is the square root of the sum of the variances of the evaluated probability distributions of the errors. It is a combination of uncertainties due to random (here referred to as random uncertainty or process precision) and systematic (here referred to as systematic uncertainty) effects. Measurement uncertainty is calculated as described in the ISO publication *Guide to the Expression of Uncertainty in Measurement*;¹³ random uncertainty from the variance of the measurement data, systematic from the variances of the probability distributions of the systematic error components. These are referred to as Type A and Type B uncertainty components, respectively. The calibration expanded uncertainty reported on the SRM certificate is two times the square root of the sum of the variances¹⁴ of all of the identified components which contribute to the measurement

uncertainty. This would correspond to the 95% confidence interval if all of the uncertainty component probability distributions were Gaussian. Vendors and buyers of materials and services should use the same method for calculating measurement uncertainties.

The effects of the random components can be estimated directly from the measurement data. Systematic uncertainty components are harder to quantify and can arise from artifact imperfections and from the measurement process.

Statistical process control A numerical value for random uncertainty cannot be determined until the measuring system is operating in a state of statistical control and the source of variability is shown to be random in nature and stochastically stationary. When these criteria have been met the process standard deviation quantifies this random uncertainty.

Each feature on every sampled SRM pitch standard is measured at least nine times over a period of several hours. Random uncertainty is determined directly from these repeat measurement data. One pitch standard has been selected to be a Control standard to serve two purposes: representative features on the Control are measured before and after every SRM calibration for statistical process control, and the pitches of several lines on the Control have been measured on the NIST Linescale Interferometer¹⁵ to provide a traceable calibration of the UV microscope's scale. This Control standard is used as long as possible to accumulate a large history of measurements.

Calibration Uncertainty components Calibration uncertainty of this pitch standard arises from several components.

TRACEABILITY TO THE METER: The definition of the meter is the length of the path traveled by light in vacuum during the time interval of $1/299\,792\,458$ of a second. The interval of the second is defined as the duration of $9\,192\,631\,770$ periods of the ground state hyperfine transition of Cs-133.¹⁶ These pitch calibrations are traceable to the NIST Linescale Interferometer, whose laser wavelength has been compared to one whose frequency was measured in a manner traceable to the second. The wavelength in the laboratory must be corrected for the index of refraction of air.

Even though the scale of the UV microscope is a laser interferometer, it is calibrated to agree with the NIST Linescale Interferometer to remove

some potential errors and to provide traceability to the meter.

ABBÉ ERROR: A significant potential error source is the Abbé error caused by possible offset between the measurement axis (defined by the laser interferometer) and the microscope focal point in combination with parasitic angular motion of the scanning stage. The measurement axis is designed to pass through the focal point of the microscope, but this is a difficult adjustment and Abbé offset in the measurement system could be as much as 1 mm. Comparison of pitch measurements made on this apparatus and on the Linescale Interferometer compensate for errors of this type.

COSINE ERROR: The error arising from misalignment of the specimen axis with respect to the interferometer measurement axis will be removed, as this angle can be inferred from the data.

TEMPERATURE EFFECTS: *STATIC* The temperature coefficient of expansion of the quartz substrate is $0.5 \times 10^{-6}/^{\circ}\text{C}$, and the longest pitch is 1 cm. If the temperature in the user's environment differs from the temperature during calibration, the worst case error is $5 \text{ nm}/^{\circ}\text{C}$.

DYNAMIC Substrate, structure, and air temperature vary with a dominant period shorter than and unrelated to the period between successive measurements. The effects of temperature variation which could contribute to systematic uncertainty are effectively randomized and average to zero because the temperature is randomly distributed among the successive measurements of any pitch. The resulting expanded random uncertainty includes all such effects and is determined directly from the data, and its components need not be individually evaluated.

It is usually advantageous to convert potential systematic uncertainties into random ones in this way, because the systematic can be difficult to estimate but the combined effect of all of the random components is determined directly. The error variations must be uncorrelated with the measurement timing. The temperature is recorded at each measurement and later tested for correlation with the pitch measurements (in their deviation from the mean).

INTERFEROMETER RESOLUTION: The resolution of pitch measurements can be limited by the resolution of the interferometer used, called the least count or least significant bit (LSB). (The inten-

sity resolution is also limited, to 16 bits). This could lead to a systematic error of $1/2 \text{ LSB}$, even in the average of any number of repeated measurements.

In this application however, the interferometer is oversampled (*i.e.*, read more frequently than the LSB would change from scanning in a noise-free environment) and digitally filtered, and—most importantly—the measurand is dithered by ambient vibration. Now the measurement resolution is limited only by noise (the dither) and not by the interferometer, and can be increased through repeated measurements. Only 1 LSB peak to peak of dither is needed. If the dither is below the Nyquist frequency but higher in frequency than the microscope resolution (spatial cutoff frequency $\times$ scan speed), it can be removed from the data. The remaining (lower frequency) dither appears in the variance of the pitch measurements.

In metrology, a little noise is a good thing.

RANDOM EFFECTS: The combined effect of all random uncertainty components appears in the variance of the repeat measurements. This includes apparatus vibration, temperature fluctuations, instrument noise, etc. The expanded random uncertainty is expected to be less than 9 nm.

Altogether, a combined expanded uncertainty for the individually calibrated standards is expected to be less than 10 nm. A more thorough analysis of calibration uncertainty is given in the documents accompanying each pitch standard.

5. Other NIST pitch standards There are two other NIST pitch standards of interest, both optimized for scanning electron microscopes.

SRM 484 is an artifact to calibrate the magnification of SEMs. This standard is fabricated by electroplating alternate layers of nickel and gold onto a substrate of a Monel sheet metal. The plate is then diced, and the individual pieces are mounted on edge in a holder. Each sample is metallographically polished to obtain a smooth surface and to reveal gold lines. Each sample is individually measured by a metrology SEM integrated with a HeNe laser interferometer. The range of spacings is from $0.5 \mu\text{m}$ to $50 \mu\text{m}$, and the expanded uncertainties, based on the ISO format (2 sigma), are 13 nm for $0.5 \mu\text{m}$ spacings and 167 nm for $50 \mu\text{m}$ spacings.

SRM 2090¹⁷ (not yet available) is intended primarily for use in calibrating the magnification scale of an SEM over a wide range of magnifica-

tions, from less than 100× to greater than 300 000×. It contains structures in both x and y dimensions, ranging in nominal pitch from 0.2 to 3000 μm, and is useful at both high and low accelerating voltages. SRM 2090 is patterned on a silicon wafer and cleanroom packaged; thus it can be inserted in modern automated measurement systems used in the semiconductor industry.

Summary A new traceable pitch standard is soon to be available from NIST. Optimized for microscopes and micromachining tools, it has pitch scales ranging from 1 μm to 1 cm, with calibrated uncertainty less than 0.1 μm. Those units which have been individually calibrated are expected to have uncertainty less than 10 nm.

Acknowledgments The author acknowledges and thanks John Beers, William Penzes, Arie Hartman, Fred Rudder, Patricia Snoots, Ted Vorburger, and the Standard Reference Materials Program, for their valuable contributions in the development of this pitch standard.

References

- 1 P.R. Nelson, P.B. Chu, and K.S.J. Pister, "Optical Methods for Characterization of MEMS Device Motion", *Proceedings of the SPIE Symposium on Microlithography and Metrology in Micromachining*, vol. 2640, pp 53-57, Austin, Texas (1995).
- 2 *NIST Calibration Services Users Guide*, NIST Special Publication SP 250 (U.S. Government Printing Office, Washington, D.C.)
- 3 J. Potzick, "Accuracy in Integrated Circuit Dimensional Measurements," Ch 3, *Handbook of Critical Dimension Metrology and Process Control*, vol. TR52, SPIE, Bellingham, Wash., pp 120-132 (Sept. 1993).
- 4 D. Nyyssonen, and C. Kirk, "Optical microscope imaging of lines patterned in thick layers with variable edge geometry: theory," *J. Optical Soc. Am.*, vol. 5 (August 1988).
- 5 M.T. Postek, "Critical Issues in Scanning Electron Microscope Metrology," *J. Res. of the National Institute of Standards and Technology*, vol. 99, No. 5 (1994)
- 6 D. Nyyssonen and R. Larrabee, "Submicrometer Linewidth Metrology In the Optical Microscope", *J. Research of the National Bureau of Standards*, vol. 92, No. 3 (1987).
- 7 NIST Photomask Linewidth SRMs 473, 475, and 476 are available from the Office of Standard Reference Materials, NIST, EM 205, Gaithersburg, Md. 20899. Voice 301-975-6776, FAX 301-948-3730.
- 8 D. Nyyssonen, "Linewidth Measurement with an Optical Microscope: the Effect of Operating Conditions on the Image Profile," *Applied Optics*, vol. 16, No. 8 (1977).
- 9 R. Silver, J. Potzick, J. Hu, "Metrology with the Ultraviolet Scanning Transmission Microscope," *Proceedings of the SPIE Symposium on Microlithography*, vol. 2439-46, pp 437-445, Santa Clara, Calif. (1995).
- 10 License information can be obtained from Terry Lynch, NIST, Gaithersburg, Md. 20899, 301-975-2691.
The Stewart Platform structure is public domain.
- 11 C. Vezzetti, R. Varner, J. Potzick, *Antireflecting-Chromium Linewidth Standard, SRM 473, for Calibration of Optical Microscope Linewidth Measuring Systems*, NIST Special Publication SP-260-119 (1992).
- 12 *International Vocabulary of Basic and General Terms in Metrology*, 2nd. Ed., ISO, Geneva (1993).
- 13 *Guide to the Expression of Uncertainty in Measurement*, ISO, Geneva (1994).
(Both available from Global Engineering Documents, 15 Inverness Way East, Englewood, Colo. 80112. 800-624-3974)
- 14 B.N. Taylor and C.E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297 (U.S. Government Printing Office, Washington, D.C., 1994).
- 15 J. Beers and W. Penzes, *NIST Length Scale Interferometer Measurement Assurance*, NIST IR 4998 (U.S. Government Printing Office, Washington, D.C., 1992).
- 16 B.N. Taylor, *The International System of Units (SI)*, NIST Special Publication 330 (1991).
- 17 Michael T. Postek, "Scanning Electron Microscope-based Metrological Electron Microscope System and New Prototype Scanning Electron Microscope Magnification Standard," *Scanning Microscopy*, vol. 3, no. 4, pp. 1087-1099 (1989)

Height Calibration of Atomic Force Microscopes using Silicon Atomic Step Artifacts

V.W. Tsai¹, T.V. Vorburger², P. Sullivan², R. Dixon², R. Silver², E.D. Williams³, J. Schneir^{2*}

¹ Dept. of Material Engineering, University of Maryland, College Park, MD 20742

² National Institute of Standards and Technology, Gaithersburg, MD 20899

³ Dept. of Physics, University of Maryland, College Park, MD 20742

*Present Address: Tencor Instruments, Santa Clara, CA 95054

Introduction

The decreasing feature dimensions required in the semiconductor manufacturing industry are placing ever increasing demands upon metrology instruments. Atomic force microscopes (AFMs), which can have lateral resolution of 1 nm and vertical resolution of 0.1 nm or below, are being used increasingly as metrology tools in the semiconductor industry. Nanometer level accuracy with AFM metrology, however, remains challenging. The scales of an instrument must be calibrated to perform accurate pitch or height measurements, while surface roughness or feature width measurements are further complicated by the 'convolution' of the probe geometry with that of the sample. Although many commercial AFMs employ some form of position sensor (capacitance or split diode) to linearize scan motions and improve repeatability, the scales of these instruments must still be calibrated using some standard artifact. At present, no AFM standard artifacts exist which are actually calibrated using AFM because no instrument with the required accuracy has been available. To address this need for AFM calibration, the National Institute of Standards and Technology (NIST) is developing a metrology AFM that will be used to calibrate traceable AFM standards. The NIST calibrated atomic force microscope (C-AFM) uses a flexure stage driven by piezoelectric actuators to scan the sample and commercial heterodyne interferometers to measure the lateral position of the sample. The control systems use two feedback loops that read from the X and Y interferometers, respectively, and adjust the piezoelectric voltages to keep the stage at the programmed lateral position. A commercial stage containing a piezoelectric actuator and capacitance transducer is used to generate and measure Z displacements (heights). A specially designed Z-interferometer head was used to calibrate the capacitance gauge in a process-intermittent manner.

Presently, there is no commercial height artifact available in the sub-nanometer range. A promising candidate for such a standard is a specially prepared silicon sample, which has single or multiple atomic steps on the surface. In this work, we will discuss our analysis algorithms and the status of our measurements of these surfaces using the C-AFM.

Experiment

The stepped silicon surface used in this work is an n type doped silicon (111) flat wafer. The sample, 10 mm x 10 mm, is cut from the wafer. The surface is cleaned at high temperature by electron bombardment heating in an ultra high vacuum (UHV) environment. To ensure the quality of the surface, a scanning tunneling microscope (STM) is used to scan the surface after cleaning. The sample is then removed from the UHV chamber. To minimize any contamination, the sample is stored in a desiccator. The steps are measured in an ambient post native oxidation. The step features on the sample can still be observed using AFM more than one year after the sample was initially prepared. The measurement of a step height with the C-AFM uses the following procedure. The first step is the calibration of the capacitance gauge with the interferometer. A sinusoidal waveform voltage is input to drive the piezoelectric tube in the Z direction. The outputs from the capacitance

gauge (Voltage) and interferometer (position in μm) are recorded simultaneously. A second order polynomial fit yields the desired scale factor of the capacitance gauge. After calibrating the capacitance gauge, the Z-interferometer assembly is replaced with the AFM head assembly. A scan is initiated over the appropriate field of view. While the area of interest is scanned, the outputs of the capacitance gauge and interferometer are then recorded using programmable commercial analytic software. Typically, images are taken with one step centered in the field of view. Generally, the terrace width between steps on our samples, which depends on the angle of surface deviation from the (111) plane, ranges from 100 nm to 2 μm . In this work, the sample used had a terrace width of about 1 μm (figure 1). We used a 1.3 μm scan size and a 0.5 Hz scan rate. Typically, one image contains 250 scan lines, each with 500 data points. In the measurement for this work, the fast scan direction is oriented across the step. An intermittent contact mode of operation was used for these measurements. The room temperature in the C-AFM laboratory is controlled to $21^\circ\text{C} \pm 0.1^\circ\text{C}$. Two sets of measurements were made in the present work. In the first experiment, we acquired twenty capacitance gauge images of the same silicon step without changing location. For the second experiment, we acquired twenty capacitance gauge images on twenty adjacent steps along the sample. The results of these two series of measurements are discussed below.

Results and Discussion

The most appropriate method for determining the step height from the C-AFM capacitance signal data has not yet been determined. Reduction of noise by averaging and removal of systematic errors are very important considerations. In figure 2, we show an average of the twenty images from the first experiment. Systematic errors have been removed in software, as discussed more below. We are presently exploring different approaches to the analysis, but for this work we have chosen to apply three methods to our data. Two major considerations in analyzing the data are: (1) the method used for averaging, and (2) the definition of the step height, or more explicitly, the location at which a height difference is to be computed. Our first two approaches involve averaging by profiles but use different definitions of the step height. The third approach uses true areal averaging and is easily employed with alternative definitions of the step height.

In the first approach, the data is leveled by subtracting a regression line fit only through one of the terraces in each profile. This serves to remove slope in the x (fast scan) axis and also detrends any slope or drift in the y direction. The fit through one only terrace is important because a fit through the entire profile of the step introduces a extra tilt which will vary with the location of the step in the profile. From the leveled data set, an analysis region is selected for both the upper and lower terraces of the step. The step height is then defined, and calculated from the entire set of scan lines, as the difference in the average heights of these two selected regions. Although this may give the appearance of areal averaging, it is not a true three dimensional method since the data set has been detrended by profiles. By application of this analysis method to the data sets obtained from our two experiments, we obtain the results shown in figures 3 and 4. The average height obtained from the first experiment is 0.272 nm, with standard deviation 0.005 nm. The average height obtained in the second experiment is 0.273 nm, with standard deviation 0.009 nm.

In the second analysis approach, a 'transition' region around the step is excluded from each profile in the leveled data set, and regression lines are fit through the remaining upper and lower portions of each scan line. The step height is then defined, and calculated for each profile, as the difference between the extrapolations of these lines at the location of the step. An average value for the entire data set is given by the average of all of the profiles in the data set. The results of our two

experiments using this analysis method are shown in figures 3 and 4. The average height of the first experiment is 0.266 nm with standard deviation 0.006 nm. The average height of the second experiment 0.269 nm with standard deviation 0.004 nm.

In the third method, no detrending of data is performed, and a true areal average is obtained by fitting planes to the upper and lower terraces of the step. It should be noted that, although these portions of the data deviate significantly from planes, the residuals for the upper and lower terraces are very similar. Hence, they represent a systematic error that can be removed by software. Such processing was applied to the average image of figure 2. Although various definitions of the step height could be used, we have chosen to extrapolate the two planes to the location of the step in the center of the image. An example of a more detailed approach that could be used would be extrapolation of the planes to a set of points that defines the step edge. These points could then be averaged or the distribution specified. Our present choice of step height definition at a single point was made for purposes of simplicity. The results of our two experiments using this analysis method are shown in figures 3 and 4. The average height of the first experiment is 0.266 nm with standard deviation 0.006 nm. The average height of the second experiment 0.269 nm with standard deviation 0.004 nm.

The observed level of agreement between these methods helps us to clarify and place bounds on the ambiguities and uncertainties involved in defining and calculating the step heights on these samples. Our future efforts will be directed toward further characterization of systematic errors in the C-AFM system and toward further consideration of the most practical and meaningful definition of the step height for these samples.

Summary

Our experiments are the first in which Si samples, prepared in UHV with single atomic steps on the surface, have been measured under ambient conditions using a metrology AFM, and also the first in which true areal analysis of the data has been performed. The difference between our measured values and the silicon (111) lattice spacing of 0.314 nm in the underlying bulk silicon is potentially due to the morphology of the native oxide on the surface and to presently unknown systematic errors in C-AFM performance for sub-nanometer scale measurements. Our future work will focus on clarifying these issues.

Acknowledgments

This work was supported in part by the National Advanced Manufacturing Testbed (NAMT) and the National Semiconductor Metrology Program (NSMP) at NIST. We give special thanks to J. S. Villarrubia of NIST for helpful discussions about this work.

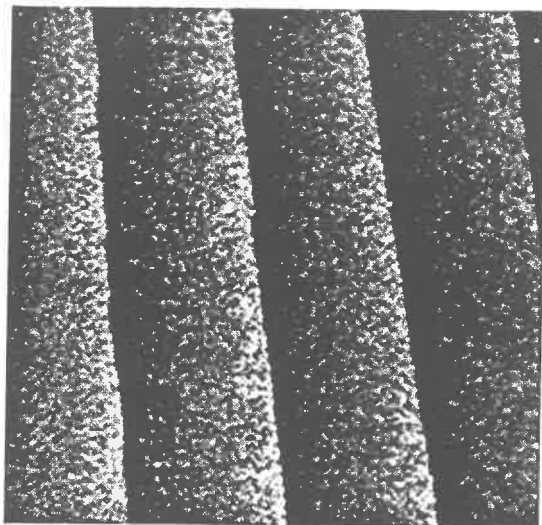


Figure 1. C-AFM image of $4.4\text{ }\mu\text{m}$ area on silicon step sample, showing the typical separation of the steps to be roughly $1\text{ }\mu\text{m}$.

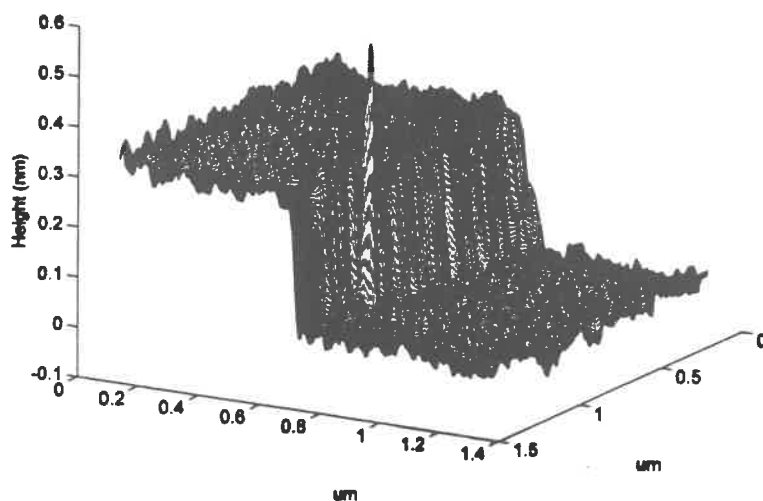


Figure 2. Average of the twenty C-AFM silicon single atomic step images, as obtained from the capacitance gauge output, with systematic errors removed in software.

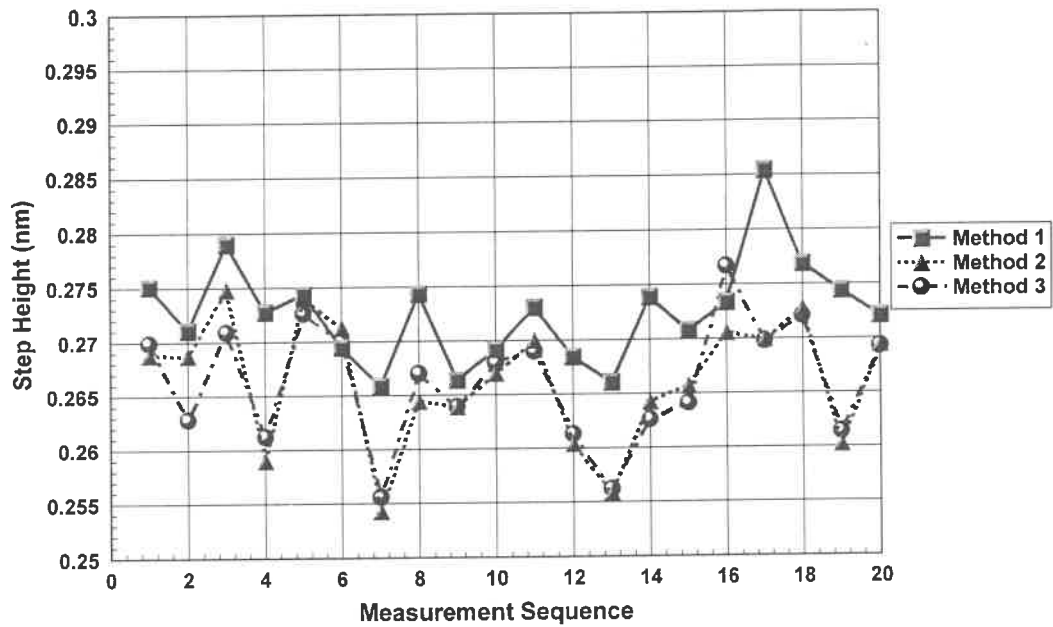


Figure 3. Step height results of the first experiment using three different analysis methods.

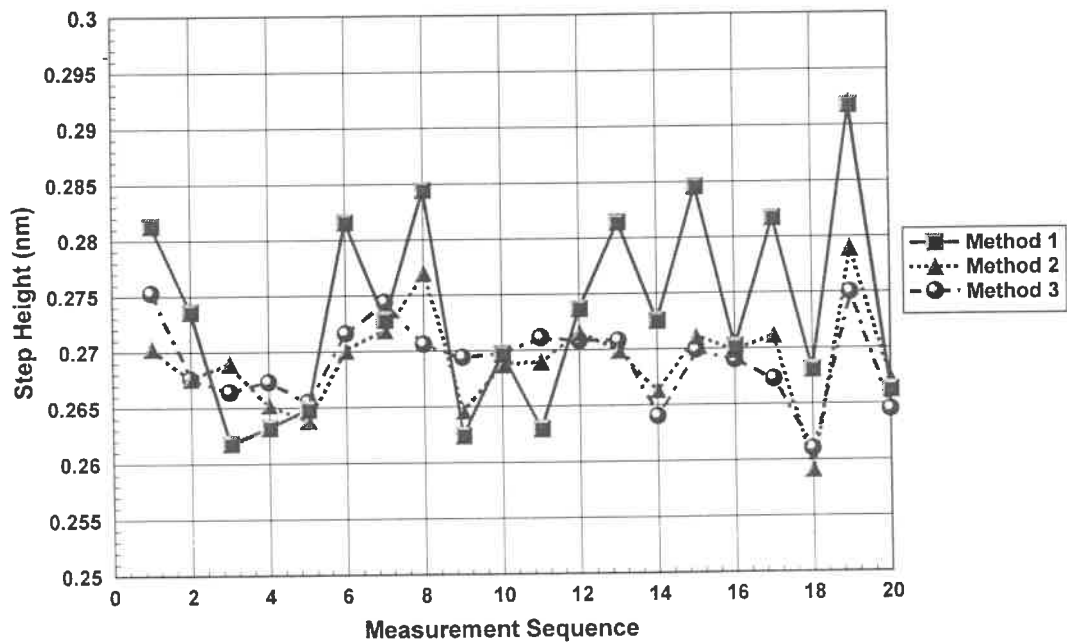


Figure 4. Step height results of the second experiment using three different analysis methods.

EVALUATING FORM ERROR OF SPHERICAL SURFACES USING CONFOCAL MICROSCOPE

Ganesha Udupa* M.Singaperumal * R.S.Sirohi @ M.P.Kothiyal @
Research Scholar Assoc. Professor Professor Professor

* Precision Engineering & Instrumentation Laboratory,
Department of Mechanical Engineering.

@ Applied Optics Laboratory, Department of Physics.
Indian Institute of Technology
Madras-600 036, India.

ABSTRACT

The surfaces obtained by various machining processes deviate from the specified geometry in terms of mathematical equations particularly regarding sphericity. Spherical surfaces pose problems in measurement and evaluation. There has been no attempt to assess a three dimensional form error on the whole surface of steel ball used for ball bearings using focus detection technique. An attempt has therefore been made to measure and evaluate spherical surfaces using confocal microscope based on focus detection technique

Key Words: Confocal Microscope, Form Error, Spherical Surface.

INTRODUCTION

Surface form error evaluation is an important aspect of precision manufacturing. The evaluation of sphericity requires a number of roundness profiles to be taken at different meridians of the sphere and assessed by fitting a reference feature on the measured data. In this paper least square sphere is used as the reference datum for evaluating the sphericity. In general, measurement of 3D form of spherical surfaces is difficult. This has been achieved with the use of CMM and talyrond roundness tester using contact principles. However this technique does not give correct assessment for form error of precision spheres due to various limitations. Optics in general played an important role in measurement and with the advent of optomechatronics is once again in the forefront of measurements. In future, it will be important to measure accurately 3D form error in order to analyse increased accuracy of balls in the lapping process and the relationship between the motion error of the bearing unit and the sphericity of the ball components of the bearing. For higher accuracy of bearings, the development of a means to measure and evaluate 3D form error of the ball using optical techniques will be important.

An optomechatronic system known as confocal microscope has been modified to measure spherical and cylindrical surfaces to an accuracy of 20 nm. In the confocal microscope, a defocused image disappears instead of blurring as in a standard microscope. This property makes it possible to image and reconstruct 3D profiles of structures with length scales of the order of a wavelength. The three basic requirements for confocal microscopes are point illumination, point detection and confocal imaging.

THE OPTICAL ROUNDNESS TESTER

An optical roundness measuring system known as confocal microscope is developed to assess the 3-D form error of spherical and cylindrical components based on the focus detection technique[1]. A focused laser spot ($\approx 0.5 \mu\text{m}$ dia) through an infinity corrected 100x, 0.90 NA microscope objective falls on the surface of the spherical part placed on a rotating table supported by an air bearing system as shown in Fig. 1. The focused laser spot scans radially with the help of a piezo electric translator (PZT). The intensity of the reflected light from the spherical surface at a point is detected by a photodetector. The position of the PZT where the intensity maximum occurs will give the relative height or the radial displacement of the spherical surface at that point. The radial displacements or the data are obtained at different positions of the sphere and the required form error is then evaluated using the least squares technique. The instrument is calibrated to an accuracy of 20 nm using a standard groove height scribed on an optical flat. The vertical resolution of the system depends on the sensitivity of the PZT used. The bearing spindle has a roundness error profile which is highly repeatable and therefore is characterized and compensated for. The traces are taken on the great circles of the sphere and sphericity is obtained from the circularity traces in space. The spindle is rotated by a computer controlled stepper motor through a set of gears which provides 400 data points to be measured around the spherical part. The bearing system with the motor is placed on a motion controlled x-y table which helps in placing the focus spot on the maximum diameter of the spherical part.

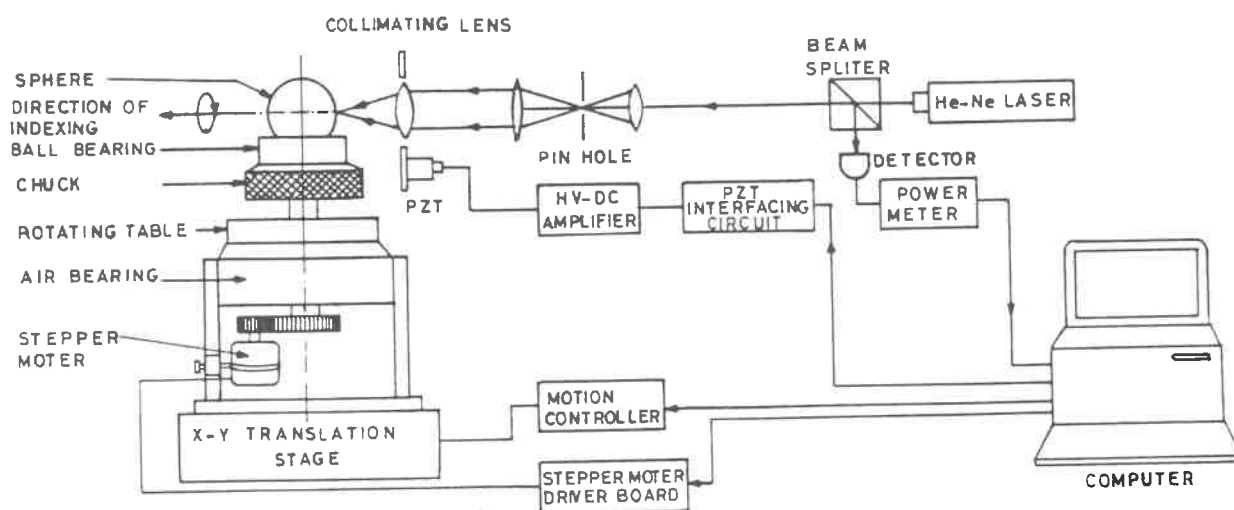


FIG.1. SCHEMATIC DIAGRAM OF CONFOCAL MICROSCOPE FOR SPHERICITY MEASUREMENT.

PRINCIPLE OF MEASUREMENT

The principle of measurement is to obtain different circularity traces by indexing the sphere and orienting the traces in space corresponding to the positions to represent the spherical surface. The traces are taken on the great circles of the sphere and the sphericity is obtained from the circularity traces in space. For this purpose the angle between the axis of the laser light and the axis of rotation of the sphere (for indexing) is 90° as shown in Fig. 1. Light small spheres are supported on a master ring gauge or on a precision ball bearing race and is indexed approximately at equal angles by visual judgment using pen marking on the sphere. Indexing of the sphere is such that it is rotated preferably, although not necessarily, at equal angles about line pq such that traces at sections 1,2 etc., are obtained (Fig.2.)

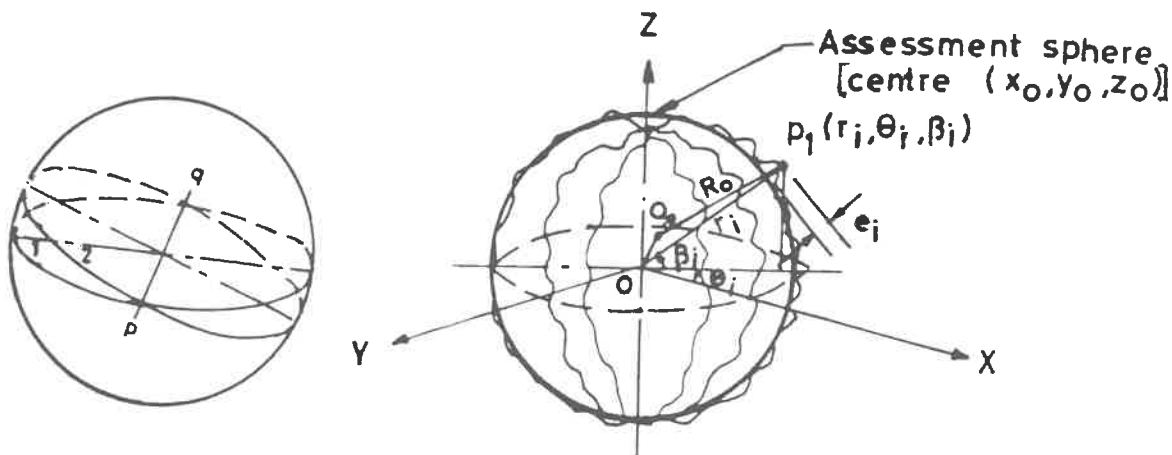


FIG.2.PRINCIPLE OF MEASUREMENT. FIG.3. ASSESSMENT OF SPHERICITY ERROR.

ASSESSMENT OF SPHERICITY

The measurement of sphericity are specified as $\{r_i, \theta_i, \beta_i\}$. If the spherical feature is well centered, then the deviations with reference to the assessment sphere (Fig.3) can be written as [2]

$$e_i = r_i - (R_0 + x_0 \cos \beta_i \cos \theta_i + y_0 \cos \beta_i \sin \theta_i + z_0 \sin \beta_i) \quad (1)$$

where $\{x_0, y_0, z_0\}$ refers to the center of the assessment sphere whose radius is R_0 .

For symmetrical data points (K observations in each meridian for θ varying from 0° to 360° , and L meridians equally spaced for β varying from $+90^\circ$ to -90°) the total observations is $KL = N$.

The least square values are

$$x_0 = 4 \sum (x_i / N) \quad (2)$$

$$y_0 = 4 \sum (y_i / N) \quad (3)$$

$$z_0 = 4 \sum (z_i / N) \quad (4)$$

$$R_0 = \sum (r_i / N) \quad (5)$$

The summation symbol in these expression represents double summation, over θ and β .

The assessment of sphericity is given by the sphericity error which is the sum of the maximum and minimum deviations from the least squares sphere.

EXPERIMENT AND RESULTS

A small steel ball of 10 mm diameter is placed on a precision ball bearing race held in a chuck and centered with the bearing axis using a micron dial indicator. The runout error of the bearing spindle is compensated using a simple reversal method as described by Donaldson [3]. The circularity trace 1 of any section on the greatest diameter of the sphere is obtained by rotating the sphere point by point and focusing the laser light at each point using the PZT. The radial height data is recorded in the computer by knowing the sensitivity and voltage supplied to the PZT in order to focus the laser spot on the spherical surface. The sphere is then indexed by a known angle such that in this position another great circle replaces the first one. Circularity trace 2 is taken in this position. The sphere is indexed six times about the line pq at 30° intervals and six circularity traces are taken and sphericity is calculated by the least squares method [2]. The traces of Fig.4 correspond to a ball of diameter 10 mm and as the points selected are equispaced the least squares sphere center is given by equations (2) - (5). The circularity error and its parameters for each trace and the sphericity error is given in the table 1.

TABLE 1 : Results of Circularity and Sphericity error

Trace No	Circularity Error μm	P - Value μm	V-Value μm	Mean line Average height μm
1	1.751	0.871	-0.880	0.378
2	1.529	0.849	-0.679	0.332
3	1.490	0.680	-0.810	0.319
4	1.061	0.479	-0.582	0.327
5	0.965	0.649	-0.316	0.285
6	1.458	0.745	-0.713	0.343
Sphericity Error = 2.820 μm				

All traces can be kept on the first trace by coinciding the pq diameters as shown in Fig 5. and the sphericity is the radial difference of two circles drawn such that they enclose all the traces with minimum radial difference. This method gave the radial difference as 3.0 μm .

CONCLUSIONS

Application of confocal microscope for the evaluation of sphericity based on focus detection technique is discussed. Small steel balls are measured with an accuracy of 20 nm. The measuring technique scans the sphere completely depending upon the number of indexes. The same technique can be extended to measure diameter and volume of the sphere. The proposed method eliminates the drawbacks of contact method.

ACKNOWLEDGMENT: The work is supported by the National Laser Programme: financial assistance is gratefully acknowledged.

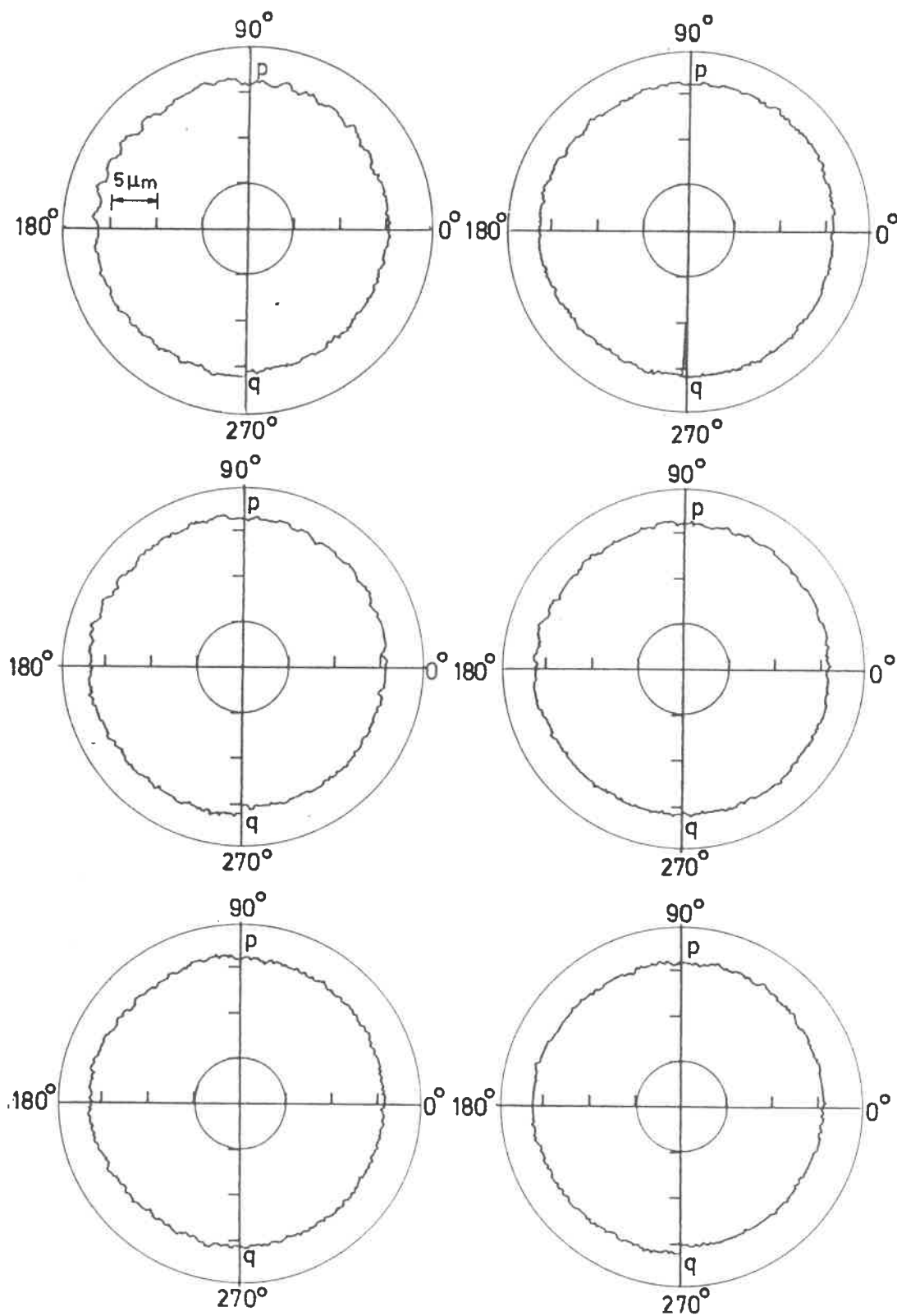


FIG. 4. TRACES TAKEN AT 30° INDEXED POSITIONS OF THE SPHERE ABOUT THE AXIS pq .

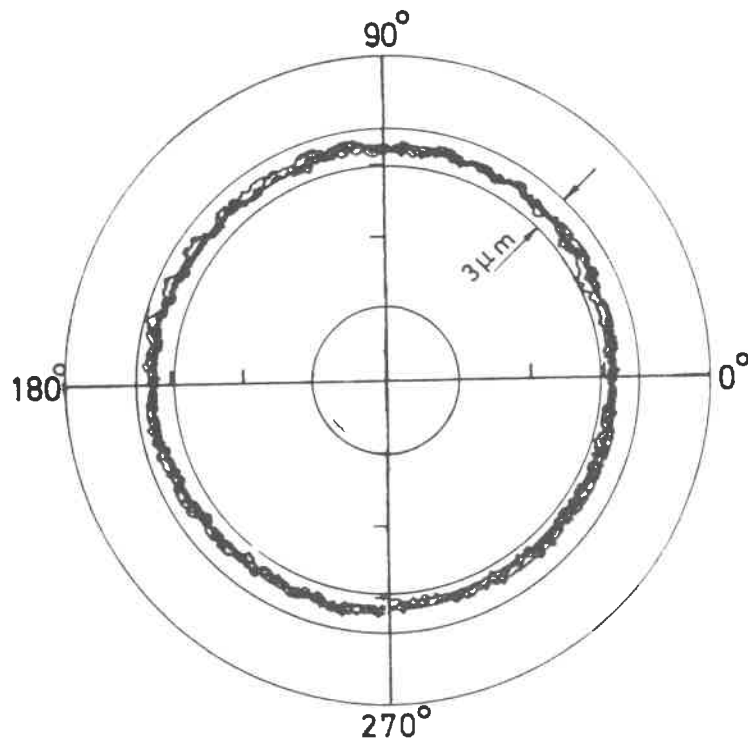


FIG. 5. TRACES ARRANGED IN A PLANE ALONG THE COMMON REPEATED POINTS WITHOUT ORIENTING THEM IN SPACE.

REFERENCES

1. Ganesha Udupa, M.Singaperumal , R.S.Sirohi and M.P.Kothiyal, Application of Confocal Scanning Optical Microscope to Three Dimensional Surface and Form Error Characterization, Proc. of ASPE Annual Meeting, Oct 15-20, 1995, pp 400-403.
2. T.S.R. Murthy, B. Raghunatha Rao and S.Z. Abdin , Evaluation of Spherical Surfaces, Wear, Vol. 57,1979, pp 167-184.
3. R.R. Donaldson, A Simple Method for Separating Spindle Error from Test Ball Roundness Error, Annals of the CIRP, V 21/2, 1972, pp 125- 126.

THERMAL ERRORS CALIBRATION AND MODELING OF A CYLINDRICAL GRINDING MACHINE CNC

DI GIACOMO, B.*; PEREIRA, P.H.**; ABACKERLI, A.J.***

* Assoc. Prof. - Universidade de São Paulo - São Carlos - SP - Brazil

** PhD Student - UNC - Charlotte - NC - USA (e-mail: phpereir@uncc.edu)

*** Assoc. Prof. - Universidade Metodista de Piracicaba - SP - Brazil

INTRODUCTION

The constant competition among industries of all countries has becoming harder in recent years due to the trend of internationalization of consumer markets. Therefore, each company interested in its own survival must continuously improve its technology to offer better products with reasonable prices.

Particularly on mechanical industries it is very important to accomplish better dimensional quality for the manufactured parts, reducing simultaneously dead time. This task can be achieved through sophisticated computer controlled machines which can produce with good quality in a short time, using practically none human intervention during the processes.

However, even such machines sometimes are not able to reach the increasing quality requirements and there are also several conventional machines being used which cannot be immediately replaced. One option to solve such problems is the compensation for the machine errors to provide more accurate working conditions.

A compensation procedure makes use of information about the machine errors that can be provided by an adequate calibration procedure of the machine individual errors. Once the error information is provided to sound a mathematical model of the machine's behavior, the influence of such errors can be evaluated and supplied to the machine control system for compensation purposes.

Among all possible machine tools error sources, the thermally induced errors are in general the most important¹. This probably is because all other error's sources have been more deeply studied for a long time, so they are better understood which makes easier their compensation.

In this context and to investigate a particular machine behavior under transient thermal influences, the objective of this work is to calibrate and model the thermal error's behavior of a CNC cylindrical grinding machine during its warmup period.

To do so, the calibration of individual errors of that machine is performed using an electronic level and an interferometer laser. Simultaneously the temperature of important thermal locations on the machine structure is monitored in such a way to create a reasonable data set to correlate thermal influences with the machine's errors. The equations for each individual error as functions of temperatures and positions are obtained by regression analysis on the error and temperature data. These equations are inserted on the model that is then completed and able to calculate the total error of the machine.

MACHINE MODELING

From all possible machine tool errors the thermal induced ones are particularly important on this work. Those deviations occurred between the tool and workpiece position due to machine tool temperature variations are considered thermal errors. This type of error is one of the most important on the final machine budget according to Janeczko, 1988 (cited by Bryan, 1990). He comments:

"As machine tool positional accuracies, tooling and cutting performance generally improve and resulting workpiece errors are reduced, thermal effects become more critical in their contribution to workpiece quality."

Besides that, the increasing cutting speeds that enlarge the heat generation and the decreasing

machine sizes that put closer the machine components, make the situation even more complicated¹³. Therefore, it is necessary to study more deeply such errors which became a major constraint for the improvement of the machine tool total accuracy.

The final error equation for the cylindrical grinding machine analyzed is obtained from a kinematic mathematical model of it. Such model considers all machine movements and their errors in order to predict the final error behavior. Each of the individual errors was calibrated along with the machine temperature evaluation and the data were analyzed. This analysis was performed to create parametric equations of each individual error as functions of the machine positions and temperatures.

Among all machine tool modeling methods the Homogeneous Transformation Technique was chosen for this work. Such choice was motivated by its simple approach for kinematic modeling, making easy to obtain the final kinematic expression for the machine. This flexibility was the main goal of F. Reuleaux, 1875 (Kinematics of Machinery, translation by A.B.W. Kennedy, Macmillan and Company Ltd., England, 1876, cited by Denavit & Hartenberg, 1955) who first tried to develop a symbolic notation to describe kinematic movements. However, according to Denavit & Hartenberg, 1955, Reuleaux's notation had some limitations and was not suitable for all types of mechanisms. They made some modifications in Reuleaux's approach and proposed what today is known as Homogeneous Transformations. Later Paul, 1981 formalized the mathematic basis about Homogeneous Transformations applied to robot manipulators, which are the same for machine tools.

Using those ideas, a kinematic model was derived for the studied grinding machine. The adopted reference frames were placed on the machine structure as shown by **Figure 1**. Each represented frame is attached to one machine element and represents its spatial position in the kinematic chain.

In the figure, the frame zero represents the absolute reference system for the whole modeling. Also, two different modeling paths can be visualized, i.e., the tool and workpiece paths. Both of them start at the origin (X_0, Y_0, Z_0) and finish at the contact point between tool and workpiece. The tool path contains the coordinate systems 0,1,2,3 and 4. In the workpiece path the frames 0,5,6,7 and 8 can be seen.

Since there are kinematic errors in both described paths, their ends in the kinematic chain are not the same, there is a difference

between them which is the total machine error. Such difference will result on workpiece errors. The modeling goal is to calculate as accurate as possible the error vector between the two modeling paths on a real situation. To do so, the total error can be calculated by a single matrix product given by the equation⁷:

$$E_{total} = T_{tool}^{-1} \times T_{work} \quad (1)$$

Where E_{total} is the total machine error matrix; T_{tool}^{-1} is the inverse transformation matrix of the tool modeling path and T_{work} is the transformation matrix of the workpiece modeling path.

The referred transformation matrices are calculated by equations (2) below which include the

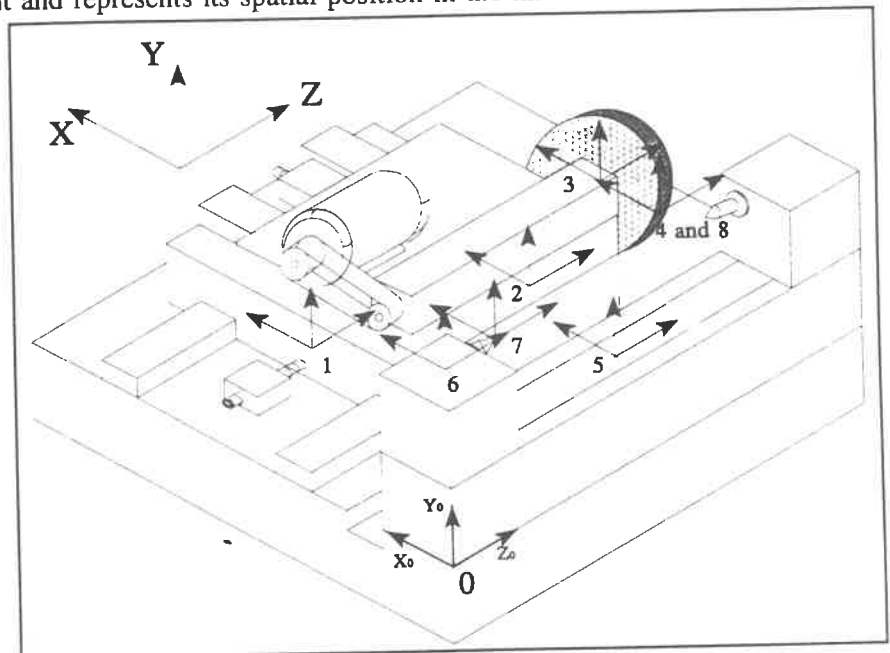


Figure 1- Coordinate systems on the machine.

individual transformation matrices of each machine element.

$$T_{tool} = {}^0T_1 \times {}^1T_2 \times {}^2T_3 \times {}^3T_4 \quad T_{work} = {}^0T_5 \times {}^5T_6 \times {}^6T_7 \times {}^7T_8 \quad (2)$$

Therefore, the final error matrix will be represented by:

$$E_{total} = \begin{bmatrix} 1 & 0 & 0 & E_x \\ 0 & 1 & 0 & E_y \\ 0 & 0 & 1 & E_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

Where the total rotation errors were not considered since in this case they have second order influence on the workpiece diameter; E_x is the resulting error on the machine X-direction. It is the most important one since it is responsible for the diameter definition, which is the main workpiece dimension; E_y will no longer be considered because its influence is of second order in the diameter; E_z has no influence in the workpiece diameter definition.

Based on these assumptions the final error equation on the X-direction is derived as follows:

$$E_x = +\delta_{xx}(x) + \delta_{xz}(z) - \epsilon_{zz}(z) \times Y_{12} + \epsilon_{yz}(z) \times Z_{23} + \epsilon_{yx}(x) \times Z_{23} \quad (4)$$

Where, to predict the final error, the individual errors below must be calibrated in order to have a complete final equation:

- $\delta_{xx}(x)$ which is the positioning error of the machine X-carriage;
- $\delta_{xz}(z)$ which is the straightness error on direction X of the Z-carriage;
- $\epsilon_{zz}(z)$ which is the angular error, "Roll", of the machine Z-carriage;
- $\epsilon_{yz}(z)$ which is the angular error, "Yaw", of the machine Z-carriage;
- $\epsilon_{yx}(x)$ which is the angular error, "Yaw", of the machine X-carriage.

MACHINE CALIBRATION AND ERROR EQUATIONS

From the model developed clearly there are five individual machine errors to be measured. From these errors only the angular error about Z of the Z-axis, $\epsilon_{zz}(z)$, cannot be measured using the laser system, instead it is used an electronic level. Both the laser and the level are used in automatic mode to provide systematic data storage during the process. In this automatic mode both errors and temperatures were recorded simultaneously by the calibration system.

Each calibration test began with the machine thermally stable. The measuring cycle started with the measurement of all temperatures, before the machine started moving by steps in forward direction. At each calibration point the error was measured and stored after the reading was stable. At the end of the course the appropriate backlash compensation was performed. Then the machine moved on the reverse direction stopping at the same points where the errors were again evaluated. After that the machine did, at the initial position and before the temperature measurement, the simulation of a small cutting cycle (without real workpiece). Before start a new cycle, another backlash compensation was again executed.

The results of the calibration of the positioning error of the machine X-carriage, $\delta_{xx}(x)$, first one to be measured. The resulting graphic is shown by **Figure 2**.

This graphic shows the measured error against the axis position where they were evaluated. Each curve represents the error behavior at a specific moment, i.e., under a different machine thermal condition. It can also be observed from this graphic that this error has a linear behavior and it is influenced by temperature variations and X-carriage position. It is also interesting to observe the small hysteresis error.

This error reaches a steady state after about nine hours of experimentation and after temperature stabilization. The temperatures reach a steady state with seven hours of experimentation as shown by **Figure 3**.

It is also possible to observe that all machine temperatures have a similar variation pattern. It can be noticed that the temperature at the spindle housing, Tp 19, increases more rapidly due to the high spindle speeds.

After the regression analysis the expression obtained to calculate the error from machine temperatures and X-carriage position is shown by equation (5) below.

$$\delta_{xx}(x) = 10^{-3} (46.210 - 0.272x - 9.320Tp_3 - 9.801Tp_4 + 8.218Tp_6 + 15.031Tp_9 - 6.660Tp_{17} + 1.264Tp_{19}) \quad (5)$$

Where:

- x is the position of X-carriage (mm);
- Tp_i is the temperature value at location "i" on the machine structure (°C).

As suggested by the discussed error variation observed on the plots this equation (5) is a linear relationship. Besides that it was executed a residue analysis to certificate that this equation agreed to the residue behavior requirements⁸. Each of the remaining errors was analyzed by the same procedure. Some of them were not so simple as equation (5), but for all errors it was possible to obtain an expression.

With all the individual error equations in the model, equation (4), it is finally possible to predict the total machine error on the X direction. Doing so, the obtained results are as shown in **Figure 4**. In this figure are illustrated

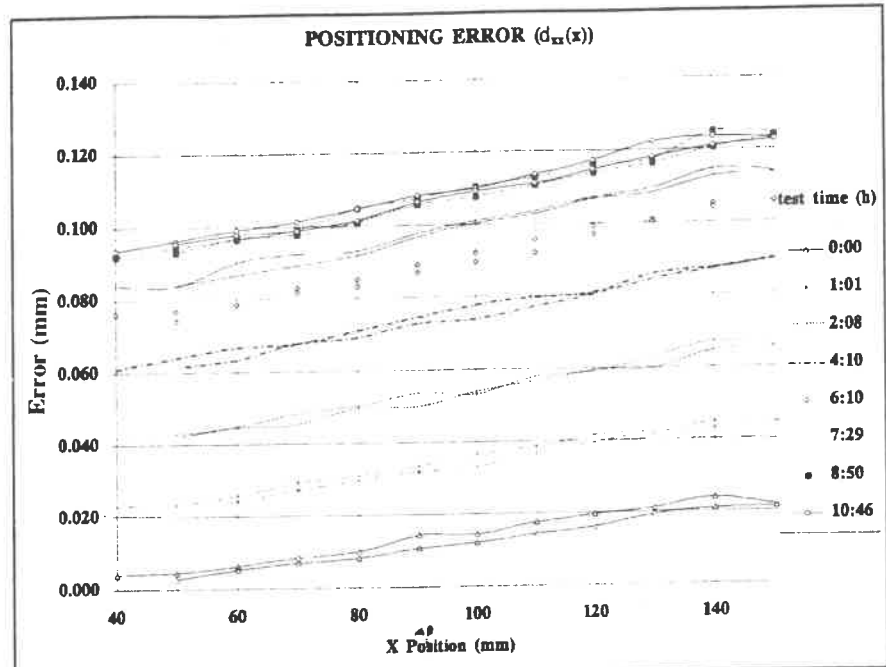


Figure 2 - Positioning error on the X-axis.

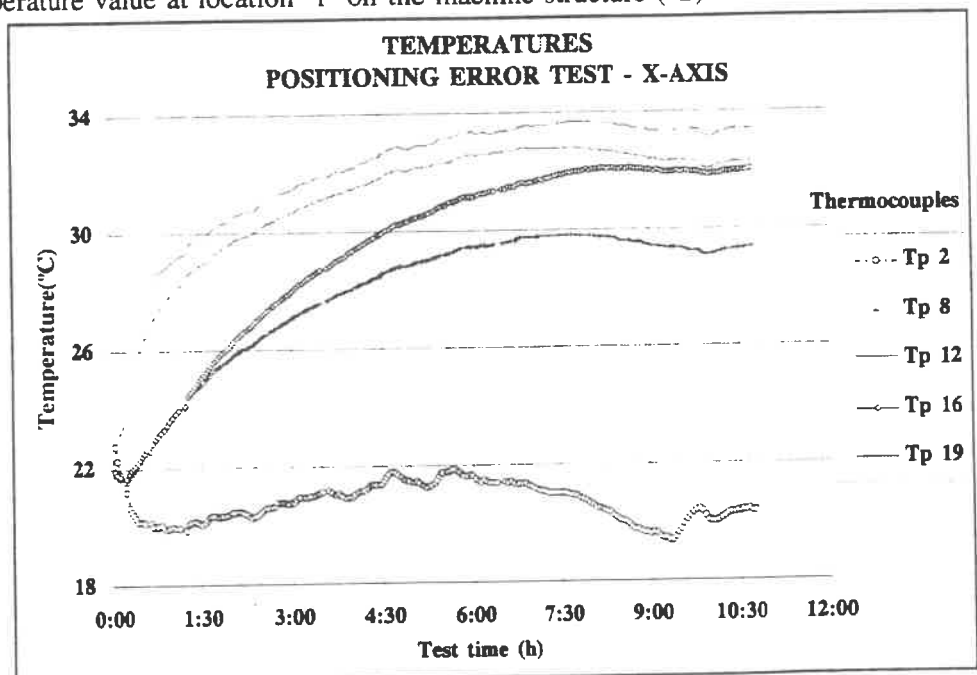


Figure 3 - Temperature variations during the calibration of $\delta_{xx}(x)$.

the error behavior in three different axis positions in terms of (X,Z) coordinates.

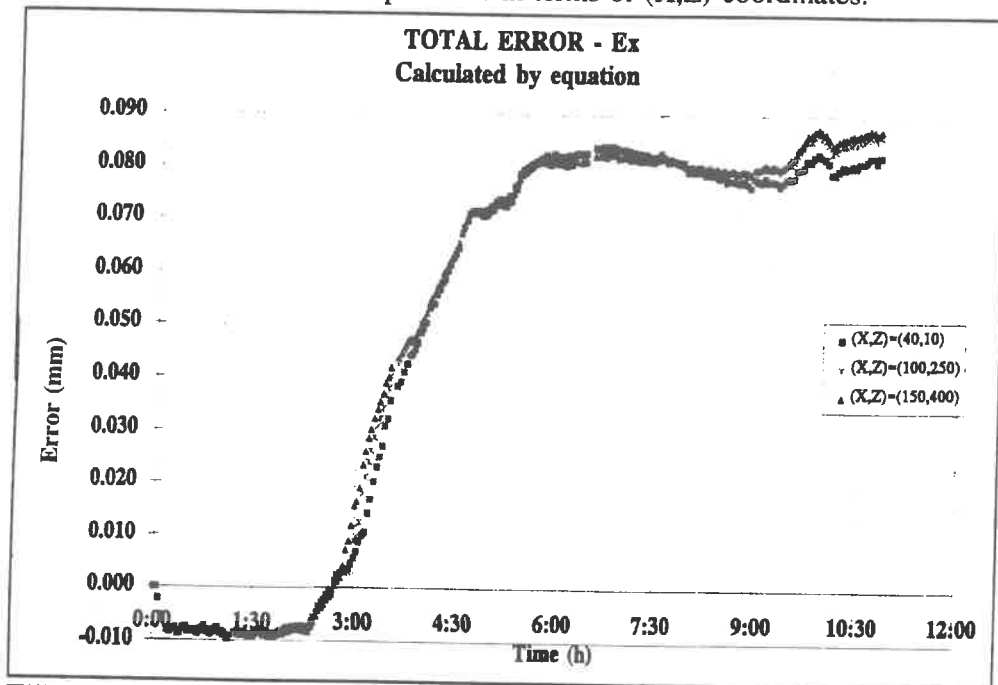


Figure 4 - Total error on direction X, E_x , calculated by the proposed equation.

One can observe that the total error at the beginning of the simulation is negative, showing that the temperature variations approach the wheel and workpiece. After six hours the error apart the wheel from the workpiece, reaching a steady state around 80 μ m gap. This might be because of the "U" shape of this machine.

It can be observed that the steady state condition is not quite stable and there is no guarantee it stays this way because the heat sources keep varying in time. If that were stable, the calculated error value could be used as a constant correction factor for the workpiece diameter after the machine has been functioning for six hours or more.

Unfortunately this is not true and therefore the temperatures must be monitored on each specific machine condition and the equation used to calculate the error values for compensation. It can also be observed that the influence of temperature variations is stronger than changes on the X, Z positions of the machine slides.

CONCLUSIONS

This work had the goal to model and calibrate the thermal error behavior of a cylindrical grinding machine mainly during its warmup period. From the work executed it is possible to point out the following conclusions:

- the temperature variations on the machine structure have a strong influence on the individual errors and also on the total one. The only exception is the Z-carriage yaw error which seems not to be influenced by the machine temperatures;
- during the initial hours of operation the total error approaches the wheel from the workpiece. After sometime this trend is inverted and the wheel is moved away from the workpiece;
- the calibration system including the temperature monitoring device developed showed to be reliable and practical. It can be used to calibrate the thermal errors of other machines;
- there are no big changes on the temperatures among close locations on the machine structure since thermocouples on such locations registered similar temperatures;
- the monitoring temperature system and the proposed final error equation can be used as a base

for a compensation system to be developed for the studied machine.

BIBLIOGRAPHY

- 01.BRYAN,J.B. *International status of thermal error research (1990)*. Annals of the CIRP, vol.39, Nº2, 1990:645-656.
- 02.ABACKERLI,A.J. *Sistema de rastreamento interferométrico para calibração volumétrica de robôs industriais*. Tese de Doutorado - EESC - USP, São Carlos, Mai.1992:1-315.
- 03.BURDEKIN,M.S. *Instrumentation and techniques for the evaluation of machine tool errors*. *Technology of Machine Tools - Machine Tool Accuracy - M.T.T.F.*, vol.5, Oct.1980:9.7.1-38.
- 04.DENAVIT,J. & HARTENBERG,R.S. *A kinematic notation for lower-pair mechanisms based on matrices*. Journal of Applied Mechanics, Jun.1955:215-221.
- 05.DI GIACOMO,B. *Computer aided calibration and hybrid compensation of geometric errors in coordinate measuring machines*. PhD Thesis - THE VICTORIA UNIVERSITY OF MANCHESTER. 1986:1-418.
- 06.DONMEZ,M.A.; BLOMQUIST,D.S.; HOCKEN,R.J.; LIU,C.R. & BARASH,M.M. *A general methodology for machine tool accuracy enhancement by error compensation*. Precision Engineering, vol.8, Nº4, Oct.1986:187-196.
- 07.DONMEZ,M.A., LIU,C.R. & BARASH,M.M. *A generalized mathematical model for machine tool errors*. Modeling, Sensing, and Control of Manufacturing Processes - 1986 ASME Winter Annual Meeting, PED vol.23/DSC vol.4, 1988.
- 08.DRAPER,N.R. & SMITH, H. *Applied regression analysis*. John Wiley & Sons, Inc. New York. 1966:1-407.
- 09.HOCKEN,R.J. *Technology of Machine Tools Volume 5: Machine tool accuracy*, Lawrence Livermore Laboratory, Livermore, CA, Oct.1980.
- 10.JANECZKO,J. *Machine tool thermal distortion compensation*. 4th Biennial International Machine Tool Technology Conference, Sep.1988:167-181.
- 11.LOVE,W.J. & SCARR,A.J. *The determination of the volumetric accuracy of multi axis machines*. Proceedings of 14th MTDR Conference, 1973:307-315.
- 12.McCLURE,R.; WECK,M. & PETUELLI,G. *Thermally induced errors*. *Technology of Machine Tools Volume 5: Machine Tool Accuracy*, Oct.1980:9.6.1-23.
- 13.MURTHY,R.L. *Thermal deformation on a semi-automatic machine: a case study*. Precision Engineering, vol.2, Nº1, 1980:17-21.
- 14.NIKITINA,I.P. & SHAKHNOVSKII,S.S. *Thermal deformations of double-sided face grinding machines*. Stanki i Instrument, vol.63, Nº7, 1992:14-16.
- 15.PAUL,R.P. *Robot manipulators: mathematics, programming, and control*. Massachusetts, The MIT Press, 1981:1-278.
- 16.PEREIRA,P.H. *Levantamento e modelamento dos erros térmicos de uma retificadora cilíndrica CNC*. Dissertação de Mestrado, USP - São Carlos, Dec.1995:1-156.
- 17.PIVOVAR,L.E. & KILOVATYI,B.Y. *How thermal deformations affect the efficiency of multispindle automatic lathes*. Stanki i Instrument, vol.61, Nº6, 1990:12-14.
- 18.SCHULTSCHIK,R. *The components of the volumetric accuracy*. Annals of the CIRP, vol.2, 1977:223-228.
- 19.TLUSTY,J. & KOENIGSBERGER,F. *New concepts of machine tool accuracy*. Annals of the CIRP, vol.19, 1971:261-273.

Substitute Geometry Fitted from Position and Surface Normal Information

Robert Edgeworth and Robert G. Wilhelm
Precision Engineering Laboratory
Department of Mechanical Engineering and Engineering Science
The University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223-0001

Introduction

Recently there have been several studies published on low cost contact and non-contact Coordinate Measuring Machine probes that are capable of measuring both the position and surface normal of inspection points [1 - 3]. Measurement of surface normals is also available on scanning CMMs. In contrast to analog scanning probes which require expensive CMM platforms to maintain their accuracy, the newly proposed probes could be used in a touch trigger mode on current, low cost probing CMMs. The simple structure of these new probes suggest that they could be produced in volume for costs which are competitive with current kinematic touch trigger probes.

Widespread implementation and acceptance of these next generation probes will depend upon the existence of applications for surface normal measurements that offer significant enhancements to the capabilities of existing coordinate measurement systems. Beyond the obvious advantages of automatically translating center of ball coordinates to the point of surface contact, little has been written about applications of surface normal measurements to coordinate metrology. The only relevant articles the authors have located are a series of articles Aoyama has published regarding feature recognition and reverse engineering of surfaces using surface measuring probes [4, 5].

The primary advantage of a surface normal measuring probe can be expressed in terms of information content. Since each sample generates six discrete pieces of data (x, y, z, i, j, k), the information generated at each point is double that collected with a standard touch trigger probe. The most obvious application of this additional information is fitting of substitute geometries to sample points. Such an approach would appear to be particularly applicable to sculptured surfaces but it could also improve 2D fitting algorithms by enhancing the stability of the least-squares fitting algorithm in the presence of form errors [9]. In this paper we study the effect of fitting to circular geometries using both position and surface normal data in the presence of systematic form errors.

Development of Fitting Algorithm

Following the method described by Forbes [6], the least squares solution for a circle is computed by iteratively solving the overdetermined matrix system:

$$Jx = -d \quad (1)$$

Where d is a $1 \times N$ vector containing the orthogonal distance to the estimated circle at each measurement point; The vector x is a 1×3 vector containing the three circle parameters (x, y, R) and J is a $3 \times N$ Jacobian matrix containing the partial derivatives of the orthogonal distance function evaluated at the sample points. The computed values of x, y and R are used to increment the initial estimate of the circle parameters and the process is repeated until the convergence criteria are satisfied.

To fit using the position and surface normal data we must introduce a second minimization function which describes the conformance of the measured surface normals to the nominal surface normals at the measurement point. Since the results of both functions will be minimized simultaneously, we must also introduce a weighting factor to control the relative influence of each function upon the fit.

$$F_1^2 + wF_2^2 = \text{minimum} \quad (2)$$

In our exploration of the effect of fitting we have intentionally selected minimization functions which will produce values of the same magnitude. For the purpose of this study we will concentrate upon fitting with a weighting factor of 1.0 but note that different weights can alter the final result.

The two minimization functions in terms of the measured data (M_x, M_y, M_i, M_j) and circle parameters (C_x, C_y, C_r) are:

$$F_1 = \sqrt{(M_x - C_x)^2 + (M_y - C_y)^2} - C_r \quad (3)$$

$$F_2 = M_i(M_y - C_y) - M_j(M_x - C_x)$$

F_1 is the orthogonal distance function specified by Forbes and F_2 is the perpendicular distance from the estimated circle center to the line through the measurement point and parallel to the measured surface normal vector. The matrix solution described in (1) now consists of a $3 \times 2N$ Jacobian matrix, a 1×3 parameter vector and a $1 \times 2N$ distance vector.

Simulated Performance Testing of Fitting Algorithm

The fitting model was first tested using data sets containing data representing perfect circles of various centers, diameters and arc lengths. In all cases the method described above performed correctly and converged quickly.

The true test of any fitting algorithm is its performance in the presence of poorly behaved data. We wished to assess the performance of least-squares fitting with position and surface normal information when the data exhibits systematic errors. To generate data with systematic form errors we used a model for generating lobed circles as described by Wilson [7]. The model generates symmetrically lobed circles by modifying the radius parameter of the circle with a sinusoidal function.

$$R_{\text{lobed}} = R - a \sin(L \theta) \quad (4)$$

The symmetry exhibited by this model is not normally encountered in actual practice but simplifies evaluation of the fitting results since the ideal center and radius parameters are unchanged by the lobing for $L > 1$.

To explore the effect of fitting with data from lobed circles, a computer simulation using the formulae described above was developed. There are five parameters which could impact the fitting solution; They are the circle radius (R), the number of lobes (L), the size of the lobes (a), the number of points sampled on the lobed circle (N), and the orientation of the sample points to the lobes (ϕ). The table below shows the range of values that were studied for each parameter.

Table 1: Parameter Ranges for Lobed Circle Fitting Study

Parameter Name	Symbol	Range
Circle Radius	R	20, 30
Lobing Amplitude	a	0.05, 0.1
Number of Lobes	L	2 - 7
Number of Sample Points	N	3 - 9
Sampling Start Angle	ϕ	0 - $2\pi/n$, 1° increments

In all cases the sample points were uniformly distributed about the circle perimeter. The least-squares circle parameters in each case were computed using both the standard, position based fitting model and the position and surface normal method developed above.

Results of Fitting Tests

The calculated radius in each individual case was the same for both fitting models. In most cases the calculated radius was correct. An incorrect circle diameter was calculated whenever the number of sample points was equal to the number of lobes on the circle. The amount of the error ranged from $-a$ to $+a$. One other case of incorrect diameters appeared whenever 3 sample points were taken on a six lobed circle. Since this was the only case in the study of a sample size that evenly divided the number lobes on the circle, it is assumed that this describes a second general rule for diameter errors. Mathematically if,

$$N = L/I \quad \text{where } I = 1, 2, 3, \dots \quad (5)$$

We hypothesize that there will be an error in the reported radius up to the amount of the lobe size. It should be noted that in all the cases described above, the correct circle center position was determined.

When the standard least squares fit computed an incorrect location for the circle center coordinates, the magnitude of the 2 norm distance from the calculated center to the true center was constant. The position errors all lay on a circle with radius equal to the value of the lobing error (a). The angular position of the center coordinates on this circle were dependent on the orientation of the sample points with respect to the lobing pattern. Similar behavior was noted by Phillips [8] in the case of fitting a circle to 3 sample points from an elliptical shape.

The position errors occurred under several different sampling conditions. It was present for all cases where the number of lobes were one greater or one less than the number of sample points. Three other cases were observed where the calculated position was incorrect. They were 5 and 7 lobed circles with three sample points and 7 lobed circles with 4 sample points. This suggests that there is another general rule for position errors which has not been identified yet.

Position errors occurred for the fit using normal vectors for all the cases described above in the least-squares fitting analysis. In contrast to the standard least-squares fits, the magnitude of the position error varies depending upon the number of samples and number of lobes. For a given sample size and number of lobes the magnitude is a constant value dependent upon the magnitude of the lobes (a). Table 2 shows the size of the errors relative to the lobe size parameter for a given sample size and number of lobes.

Table 2: Magnitudes of Position Errors for LJK Fit

Sample Size	Number of Lobes	Position Error
3	2	0.5a
3	4	2.5a
3	5	2a
3	7	4a
4	3	a
4	5	3a
4	7	3a
5	4	1.5a
5	6	3.5a
6	5	2a
6	7	4a
7	6	2.5a
8	7	3a

Note that only for one case, 3 sample points on a 2 lobed circle, is the position error less for the fitting with surface normals than it is for the standard least-squares fits. Reducing the weighting of the surface normal data in the fit increases the number of cases which perform better than a standard least-squares fit, but no value of the weighting resulted in fits that were better than the standard estimator for all lobing cases. Figure 1 shows the geometry involved for two of the fitting cases that resulted in errors.

Conclusions

We have developed a least-squares circle fitting algorithm which employs both position and surface normal measurements to compute the circle parameters. Testing reveals that, in the presence of symmetrically lobed measurements, using surface normals to fit substitute geometry can magnify the errors encountered using a standard least-squares fit and in only limited cases are the results improved in comparison to the standard fit. However, the degenerate fitting cases determined in this study are detectable by the feature form when using the surface normal fit, in contrast to the standard fit, which reports perfect form and therefore gives no indication that the fit has failed. This suggests that, with selection of an optimum weighting factor, fitting with surface normal data could enhance the performance of Coordinate Measurement Systems. Investigation of optimal weighting factors is currently underway.

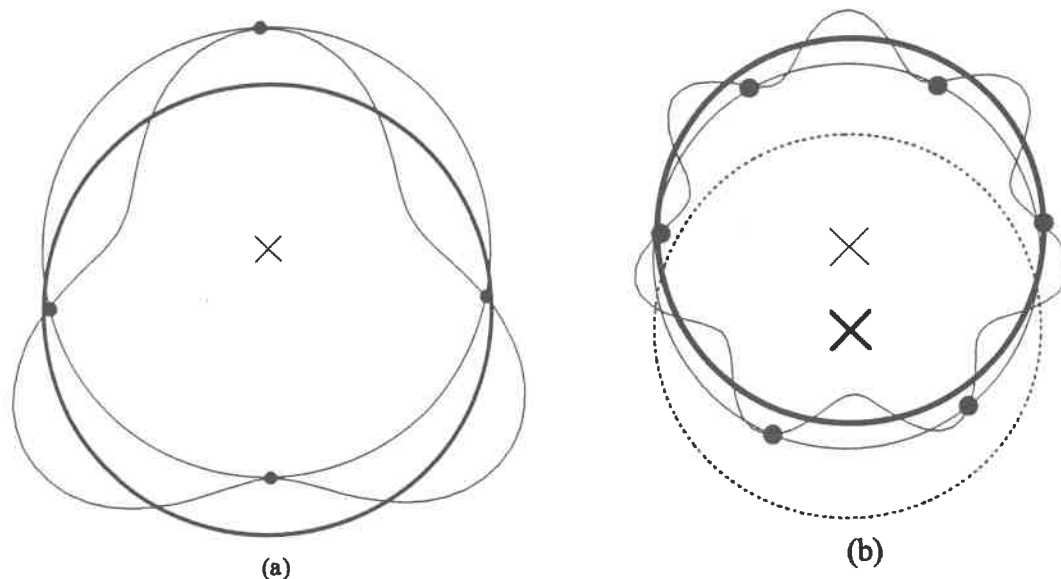


Figure 1: Examples of sample points on lobed circular geometry. (a) Four point measurement on 3 lobed circle. The bold circle represents the nominal shape, the lighter circle and marker is the substitute geometry calculated by both types of least-squares fits. (b) Six point measurement on 7 lobed circle. The dotted circle represents the circle parameters found using position and surface normals and the light solid line is the result of the standard least squares fit.

References

1. Aoyama, H., Kawai, M. and Kishinami, T., "A New Method for Detecting the Contact Point between a Touch Probe and a Surface", *Annals of the CIRP*, 1989, 34/1, 593 - 596
2. Yamazaki, K., Aoyama, H., Lee, K. S., and Sawabe, M., "Non-Contact Probe for Continuous Measurement of Surface Inclination and Position using Dynamic Irradiation of Light Beam", *Annals of CIRP*, 42/1, 585-588.
3. Vidic, M., Harb, S. M., and Smith, S. T., "Observations of Contact Measurements Using a Resonance Based Touch Sensor", *Proceedings of the American Society of Precision Engineers*, 1996
4. Aoyama, H. and Inasaki, I., "Surface Classification and Model Construction Based on Processing Position Vectors and Unit Normal Vectors", *Transactions of NAMRI/SME*, 24, (1996), 211-216
5. Aoyama, H., et al., "Measuring Method of Shape Elements by Detecting Position Vectors and Unit Normal Vectors", *Proceedings of Japan-USA Symposium on Flexible Automation*, 1994, 563-568
6. Forbes, A. B., "Least-Squares Best-Fit Geometric Elements", NPL Report DITC 140/89, April 1989
7. Wilson H. R. Jr., "Form Errors of Cylindrical Features for Specific Manufacturing Processes", Master's Thesis, University of North Carolina at Charlotte, 1992
8. Phillips, S. D., "Chapter 7: Performance Evaluations", *Coordinate Measuring Machines and Systems*, J. A. Bosch, Ed., Marcel Dekker Inc., New York, 1995
9. Hocken, R. J., Raja, J., Babu, U., "Sampling Issues in Coordinate Metrology", *Manufacturing Review*, 6/4, Dec. 1993

Modeling and Simulation of a Parallel Architecture, Five-axis Coordinate Measuring Machine

Bernhard Jokiel, Jr. and John Ziegert
University of Florida, Gainesville

Introduction

With the growing interest in machine tools based on parallel linkage architectures, it has become important to develop a coordinate measuring machine (CMM) better suited to accurately measure the complex, contoured parts these machines can produce. A CMM based on parallel linkages was introduced by Jokiel and Ziegert (2,3) to provide measurement in five degrees-of-freedom (DOF) (X, Y, Z, θ, ϕ). This paper presents results from the simulation of such a CMM, including machine sizing, workvolume determination, forward and reverse kinematics, required actuation forces and predicted accuracy.

Hexahedral CMM revisited

Five actuated (legs 1,2,4,5,6) and one passive (leg 3) prismatic, telescoping struts are arranged in a hexahedral configuration (Figure 1). The sphere-center to sphere-center length of each strut is sensed by a laser interferometry system. Knowing the strut lengths, the distance between the base spheres, and the probe length, the position and orientation of the probe tip can be determined.

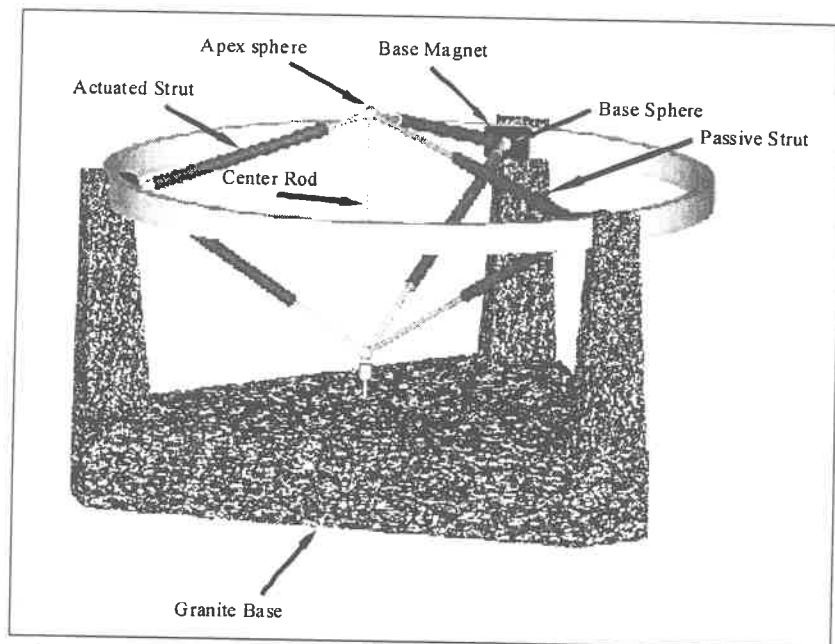


Figure 1 - Hexahedral CMM

Machine sizing

The workspace of the hexahedral CMM is not a simple, easily recognizable configuration (like a rectangular box for bridge-type CMM's), and relies heavily upon realizing the proper dimensions of the strut components. By assuming values for the dimensions of a simplified, cylindrical workvolume, the minimum and maximum strut lengths, and base and apex sphere diameters, the ideal dimensions of the machine components (strut extension length, center rod, and the distance between the base spheres - Table 1), and machine limits (strut and center rod angular interference) may be determined by an iterative process.

Assumed Dimensions	inches (m)	Calculated Dimensions	inches (m)
Minimum strut length	30.0 (0.762)	Maximum strut length	50.7 (1.288)
Workspace diameter	15.0 (0.381)	Machine diameter	71.51 (1.816)
Workspace height	15.0 (0.381)	Center rod length	35.15 (0.893)
Center rod sphere dia.	3.00 (0.0762)	Base triangle length	61.93 (1.573)
Base sphere dia.	3.00 (0.0762)		

Workvolume boundary

Two methods were used to determine the boundaries of the useable workvolume. The first determines the mobility plane (Figure 2) which is the area where a vertically oriented probe can move, constrained at a constant Z, without violating the minimum or maximum strut lengths, or strut interference angles. The second determines the mobility cone (Figure 3). The boundary of the mobility cone determines the maximum probe tilt (ϕ), for the probe rotation (θ), at a fixed probe tip coordinate (X_P , Y_P , Z_P), without violating the minimum or maximum strut lengths, or strut interference angles.

Kinematic transformations

The reverse and forward kinematics have closed-form solutions. The forward position of the probe tip is expressed in terms of variables X_{S2} , Y_{S2} , Z_{S2} (coordinates of the lower center rod sphere), θ (probe tilt), and ϕ (probe rotation), all of which are ultimately expressed in terms of the six strut lengths (L_i $i=1$ to 6) and the base lengths (LB_1 , LB_2 , LB_3) and the probe length (L). By taking the time-derivatives of these relations, expressions for the forward velocity, acceleration and jerk may be obtained (not shown).

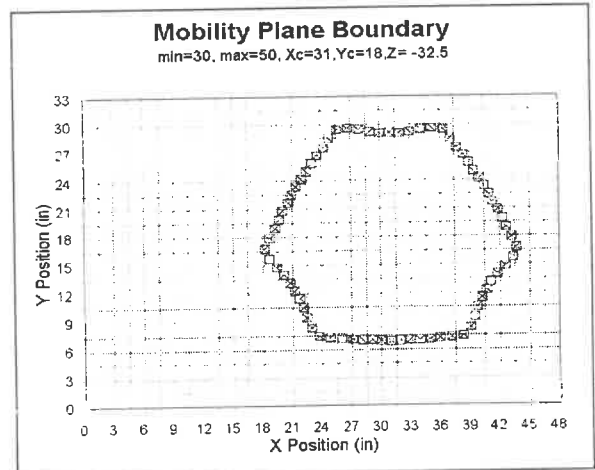


Figure 2 - Mobility plane example.

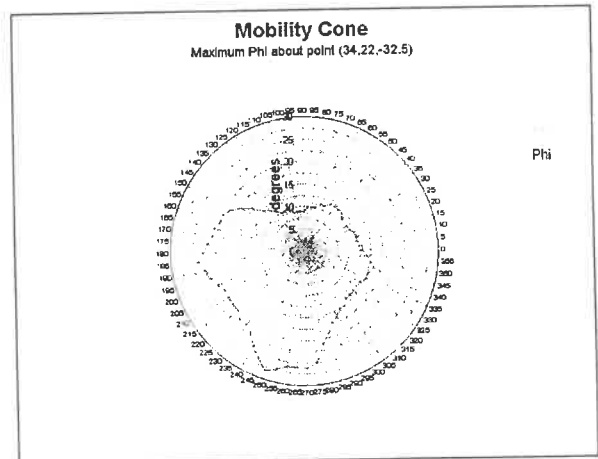


Figure 3 - Mobility cone example.

$$\begin{aligned}
 b &= LB_2 \left(\frac{LB_1^2 - LB_2^2 - LB_3^2}{-LB_2 LB_3} \right) & h &= +\sqrt{LB_2^2 - b^2} \\
 X_{S_1} &= \frac{1}{2LB_2} (L_1^2 - L_3^2 + LB_2^2) & X_{S_2} &= \frac{1}{2LB_2} (L_2^2 - L_4^2 + LB_2^2) \\
 Y_{S_1} &= \frac{1}{2h} (L_1^2 - L_5^2 + b^2 + h^2 - 2bX_{S_1}) & Y_{S_2} &= \frac{1}{2h} (L_2^2 - L_6^2 + b^2 + h^2 - 2bX_{S_2}) \\
 Z_{S_1} &= +\sqrt{L_2^2 - X_{S_1}^2 - Y_{S_1}^2} & Z_{S_2} &= +\sqrt{L_2^2 - X_{S_2}^2 - Y_{S_2}^2}
 \end{aligned}$$

$$L_c = +\sqrt{(X_{S_1} - X_{S_2})^2 - (Y_{S_1} - Y_{S_2})^2 - (Z_{S_1} - Z_{S_2})^2}$$

$$\phi = \cos^{-1}\left(\frac{Z_{S_1} - Z_{S_2}}{L_c}\right) \quad \theta = \tan^{-1}\left(\frac{Y_{S_1} - Y_{S_2}}{X_{S_1} - X_{S_2}}\right)$$

$$X_P = X_{S_2} - L_P \sin \phi \cos \theta \quad Y_P = Y_{S_2} - L_P \sin \phi \sin \theta \quad Z_P = Z_{S_2} - L_P \cos \phi$$

Strut actuation forces

In order to determine the proper type and size of the actuators required, it is necessary to determine the static and dynamic actuation forces required in each leg. The upper tetrahedron legs were analyzed first, then the center rod, and finally the lower tetrahedron legs. This analysis is too long to be presented here. The largest component of the total actuation force (dynamic + static) was the static part. The maximum force is located in the lower tetrahedron legs, since they must support the weight of the center rod and the upper tetrahedron legs. The actuation for in the lower legs approaches infinity as they retract toward the plane of the base spheres (indicating that this plane is a singularity in the workspace). Figures 4 and 5 show the variation of the maximum actuation static force. Figure 4 is a slice through the top of the workvolume, and 5 is a slice through the bottom of the workvolume. Figure 6 shows the variation of the actuation force during a linear movement.

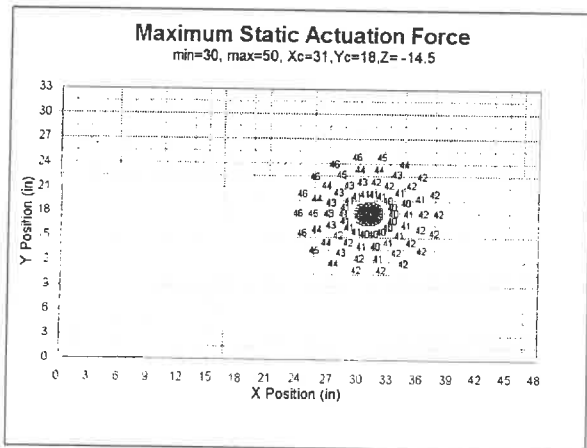


Figure 4 - Maximum static forces at the top of the workspace.

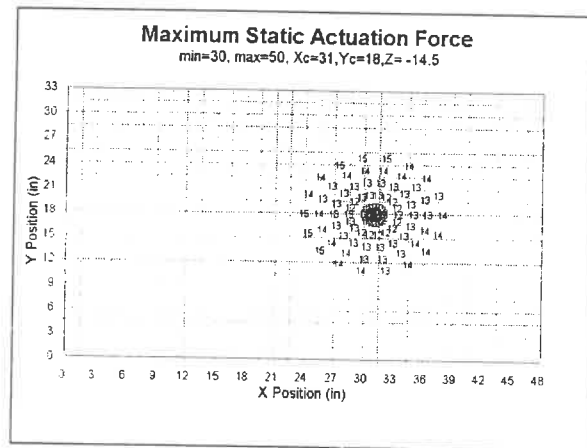


Figure 5 - Maximum static forces at the bottom of the workspace.

Predicted accuracy

The accuracy of the CMM was determined using the forward equations of motion and a multivariate perturbation technique during a static machine state. The first step was estimating the error present in the sphere-center to sphere-center measurement of each strut. This was achieved by an error budget. Some of the sources of error considered were: thermal growth of unsensed lengths, cosine alignment of laser deadpath, glass path error, beam misalignment due to

rotation of optics, and bending of the strut. Using the multivariate perturbation technique from Beasley and Figliola (1), and the values from the error budget for the deviation in the lengths of the six struts, and distances between the base spheres, the approximate error in the position of the probe tip could be determined. The error in the length measurement of each strut was $\pm 41\mu\text{in}$ ($\pm 1\mu\text{m}$). The total error at the probe tip varied from $\pm 96\mu\text{in}$ ($2.4\mu\text{m}$) to $\pm 213\mu\text{in}$ ($5\mu\text{m}$).

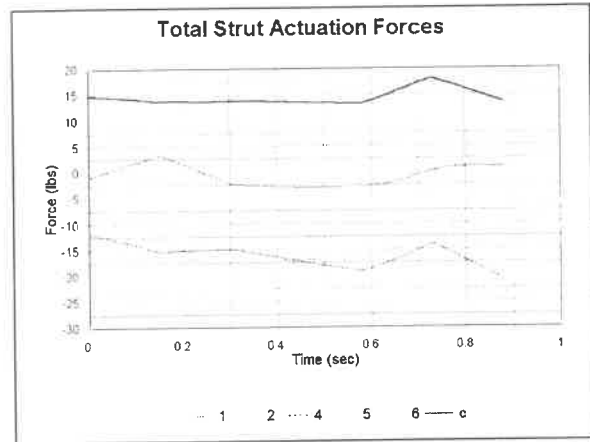


Figure 6 - Total required strut actuation forces during a linear movement.

Conclusion

The hexahedral CMM concept is a viable solution for the high precision measurement of complex, contoured workpieces, and offers a new alternative to serial linkage, bridge-type CMM designs. It offers a completely kinematic frame design, with laser feedback measuring system capable of locating a probe in five DOF with a high degree of accuracy. Since most of the strut length is sensed by the laser system, is more robust to error induced by thermal effects. Two interesting possibilities occur with the fixed length center rod which have not been discussed in this paper. One is that the machine is self calibrating and the second is that it is potentially self-adjusting to thermal disturbances occurring in the machine frame. These two subjects are currently under investigation. Although the concept is good, realizing a machine will be difficult. The hexahedral CMM does have a small workvolume for its footprint. Some of the design challenges which the Machine Tool Research Center (MTRC) at the University of Florida face are designing struts which have longer real extension lengths, and joints which are stiffer, collapse to narrower angles, while minimizing the amount of strut length unsensed by the laser feedback system.

Works cited

1. Beasley, D.E., Figliola, R.S., Theory and Design for Mechanical Measurements, John Wiley & Sons, New York, pp. 125-209, 1995.
2. Jokiel, B., Ziegert, J.C., Development of a Parallel Architecture, Five Axis Coordinate Measuring Machine, ASPE Annual Meeting Proceedings, Austin, Texas, pp. 313-316, 1995.
3. Jokiel, B., Analysis of a Parallel Architecture Coordinate Measuring Machine, Master's thesis, University of Florida 1996.
4. Mize, C.D., Design and Implementation of a Laser Ball Bar Based Measurement Technique for Machine Tool Calibration, Masters thesis, University of Florida, 1993.

EVALUATION OF EFFECT OF SAMPLE POINT DENSITY ON DIMENSIONAL MEASUREMENTS OF CYLINDRICAL FEATURES USING COORDINATE MEASURING MACHINE

A. Chukwujekwu Okafor and Deepak Mehrotra
Laboratory for Industrial Automation and Flexible Manufacturing
Department of Mechanical and Aerospace Engineering and Engineering Mechanics
University of Missouri-Rolla
Rolla, MO 65409-0050

ABSTRACT

The current trend towards automation and high precision of manufactured parts have necessitated the automation of manual inspection to keep pace with the increased productivity and maintain repeatability in measurement. Thus to automate the inspection process, coordinate measuring machines (CMMs) are used for dimensional gaging of mechanical parts. The use of automated CMMs for inspection and process control has resulted in the identification of technological voids that need further research and understanding. Due to the absence of an established standard for inspection, most computer assisted CMMs currently use different measurement techniques and analysis routines, which produce different results from the measurement of the same part features. Thus there is a need for an explicit mathematical definition of geometric dimensions and tolerances (e.g. ANSI Y14.5), along with an associated standard of dimensional measurement techniques and methods.

This paper is a contribution to the establishment of a standard for measuring cylindrical features using a CMM. The paper presents the results of the evaluation of the effect of sample point density on dimensional measurements of machined cylindrical features. Outside-diameter (OD) and inside-diameter (ID) cylindrical specimens were machined from 6061-T6 aluminum on a CNC lathe. The geometric entities such as roundness, tolerance and concentricity of the OD turned and ID bored parts were measured off-line using a CMM. The results show that the uncertainty, repeatability, and statistical reliability in CMM measurements of roundness, tolerance and concentricity vary with the number of sample points used. The higher the sample point above the recommended minimum values the smaller the variability, uncertainty, and the better the repeatability. The effect of sample point density on CMM measurements of these geometric entities increases with tool wear.

INTRODUCTION

To automate the inspection process, coordinate measuring machines (CMMs) are used for dimensional gaging of mechanical parts. The use of automated CMMs for inspection and process control has resulted in the identification of a number of technological voids that need further research and understanding [1]. Due to the absence of an established or agreed upon inspection techniques standard, most computer assisted CMMs currently use different measurement techniques and analysis routines, which produce different results from the measurement of the same part feature. Thus there is a need for an explicit mathematical definition of geometric dimensions and tolerances (e.g. ANSI Y14.5), along with an associated standard of dimensional measurement techniques and methods.

The first void mentioned-explicit product definition, was brought into national prominence through GIDEP (Government Industry Data Exchange Program) Alert issued by R. K. Walker of Westinghouse Corporation [2]. The alert takes issue with the clarity and effectiveness of using ANSI Y14.5 drafting standard as a "gaging" standard. Also, in the May 1990 issue of Manufacturing Engineering, in an article entitled "The Methods Divergence Dilemma," [3], it is reported that "significant roadblock has appeared in the reliable use of the current ANSI Y14.5 drafting standard for coordinate measuring machines". This roadblock as explained

is based on methods divergence - the difference in methodology between CMMs and hard gages. The first problem is now being addressed by ANSI Y14.5 ad hoc group under the auspices of ASME.

The second void mentioned - the need for the establishment of an inspection technique standard, was also first recognized by R. K. Walker at Westinghouse after an episode in late 1988 [1]. As reported, they assessed some parts on their CMMs and found their flatness and parallelism to be in opposition which Walker explained is a mathematical impossibility according to ANSI Y14.5. He then followed this on and designed a test database of part geometries and asked their CMMs which came from various manufacturers - to assess flatness, parallelism, perpendicularity and so forth. According to Walker, there were radical differences in the answers he got from the same measurement data. He ended up writing a GIDEP Alert on all the data they had put through the CMMs. The alert reads: "Most computer-assisted CMMs are using incorrect methods to calculate inspection results. Errors may be as large as 50%".

Several attempts have been made for "mathematizing" the current Y14.5 standard. Requicha [4] and Requicha and Chan [5] formulated a mathematical framework for tolerances based on the use of offset boundaries. The space between the offset boundaries defines the tolerance zones as per the theory presented by them. Although the concept of tolerance zone is useful in describing individual tolerance requirements, it suffers from many departures from the current standard.

Using virtual boundaries which act as separating surfaces between the allowable regions of space occupied by different parts in an assembly, Srinivasan et al. [6] tried to capture the requirement for non-interference between parts.

Realizing that mathematizing the tolerance is insufficient for application of CMMs to 3-dimensional metrology, Hocken [7] is studying to control the effects of measuring systems and sampling algorithms on the inspection results. The research reported in this paper is also aimed in the same direction and presents simple procedures that are applicable in many situations and sampling strategies that enable analysis of complicated surfaces and interrelationships among features.

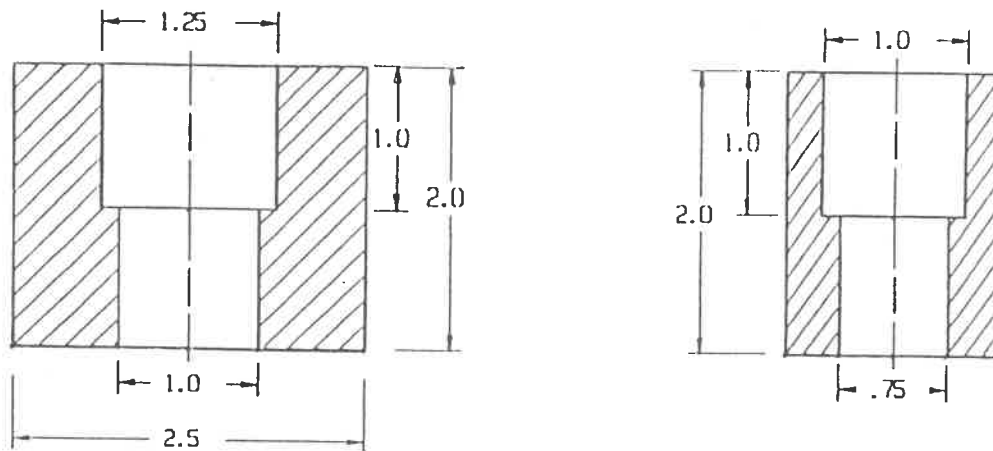
Thus much work has been done towards "mathematizing" the current Y14.5 standard for coordinate measuring machines but with limited success. The current study is a contribution towards the Computer Aided Manufacturing-International's (CAM-I) Quality Assurance Program supported efforts towards establishing dimensional inspection techniques for sample-point dimensional measuring equipment. Dimensional inspection techniques include the determination of measurement sampling, distribution of samples and selection of mathematical methods used to represent the actual geometry. These techniques are based on conditions such as feature type, geometric size, feature function, manufacturing process, tolerance aspect, tolerance limits and confidence level. In this paper the effect of sample point density on dimensional measurements of cylindrical features by CMM is presented. Theoretically, there are a minimal number of points required to determine geometric entities [8]. However, additional points are necessary to gain concise, statistically reliable information about the object's true departure from nominal. The effect of additional points are investigated in this research.

EXPERIMENTAL SET-UP AND MEASUREMENTS

Experiments were conducted on OKUMA LB15 CNC Lathe. The drawing of the machined components and their dimensions are shown in Figure 1. The design of the workpiece was decided to conform to Computer Aided Manufacturing-International (CAM-I) Inc. request for quotation for research and development of dimensional inspection techniques specification (DITS) phase I. DITS is intended to be a guideline for using the Coordinate Measuring Machines (CMMs) for the inspection and process control of cylindrical features in accordance to the design intent specified by ANSI Y14.5M-1982.

The coordinate measurement system used for evaluating the effect of sample point density on measurements of cylindrical features was a Brown and Sharpe Micro-validator CMM with an accuracy of 0.0001 in.. Form tolerances in the form of circularity or roundness, location tolerances in the form of concentricity and runout tolerances in the form of circular runout are the different geometric characteristics of cylindrical part features that were measured off-line.

The rigid ball-point probe tip of the manual CMM was qualified against a reference sphere mounted on a post that was fastened to the CMM table. It was followed by alignment, in order to compensate for differences between the coordinate system of the part and the coordinate system of the CMM. The primary datum was established at the surface of special clamp/fixture designed and machined at the machine shop. All cylindrical feature measurements were taken 0.9" below the top of the cylindrical part.



All dimensions in inches

Figure 1: Experimental workpiece drawings

To study the effect of sample point density, the measurements of cylindrical features: tolerances, roundness and concentricity were first done on one machined workpiece. Tolerances for the inside and outside diameters (ID & OD) were measured with evenly spaced sample points around the part circumference in steps of 3, 6, 9, 12 and 15 sample points respectively. Roundness or circularity of the inside and outside diameters were measured with evenly spaced sample points around the part circumference in steps of 4, 8, 12, 16 and 20 sample points respectively. Concentricity of the cylindrical part was measured with evenly spaced sample points around the part circumference in steps of 6, 12, 18, 24 and 30 sample points respectively. Concentricity is a condition in which two or more features in any combination have a common axis.

RESULTS AND DISCUSSION

Figure 2 shows the variation of OD roundness measurement with sample point density. The measurements were taken in steps of 4, 8, 12, 16 and 20 sample points respectively. At each of these steps 4 replicates were taken in order to study the repeatability and uncertainty during the measurement. Standard deviation σ was used as an estimate of machine repeatability.

It can be seen from Figure 2 that the curve of OD roundness measurement steeply rises up till 12 or 16 sample point measurements after which the curve flattens out. The measurement points seem to be more spread out at low sample points but the spread narrows down at higher sample point density. This can also be confirmed by the standard deviation values shown on the plots. This implies that the error (measurement uncertainty) in the sample point measurement decreases with the increase in number of sample points for OD roundness measurement. In other words measurement uncertainty, repeatability and statistical reliability of the sample point measurements improve with increase in the number of sample points above the recommended minimum. The measurement uncertainty is lowest at 20 sample points. This variation could be attributed to different error sources affecting the coordinate measuring machine performance. These include geometric

errors, thermal distortions, kinematic errors and probe/workpiece errors [9]. As more sample points may not be economically justified due to the expense of measurement time and computational performance, 12 or 16 sample points are recommended for OD roundness measurement. Figure 3 showing the variation of ID roundness measurement depicts similar characteristics as OD roundness except that the curve rise in the left half (low sample points) is steeper. Repeatability of the measurements improve with more sample points. As with OD and ID roundness, the measurement uncertainty is lowest at 20 sample points, but this large number of points may not be economically justifiable. Therefore 12 or 16 sample points is recommended for ID roundness measurements as well. Curves of variation of ID/OD tolerance measurements with sample point density show a slight positive trend for ID tolerance and negative trend for OD tolerance as shown in Figure 4 and 5, respectively. Measurements with 3 and 15 sample points gave the least measurement uncertainty. Based on this information 3 to 15 sample points are recommended for reasonable accuracy in the measured values. Curve showing variation of concentricity with sample point density is almost flat (see Figure 6). Repeatability for the measurement of concentricity also improves with larger sample points. 18 and 30 points gave the least measurement uncertainty and would be recommended for reliable accuracy in the measurement of concentricity.

CONCLUSIONS

1. Sample point density affect CMM measurements of tolerances of cylindrical features. This effect is more conspicuous when parts machined at the same cutting condition and the same tool are compared, which leads us to believe that same sample point measurement strategy may not be applicable to the parts not machined at same cutting conditions.
2. For cylindrical features, CMM measurements of OD/ID roundness and tolerances generally increase with an increase in sample point density. The higher the sample point above the recommended minimum values, the smaller the variability and uncertainty, and the better the repeatability and statistical reliability in the CMM measurements of roundness, tolerances and concentricity. This variation could be attributed to different error sources affecting the coordinate measuring machine performance.
3. The effect of sample point density on CMM measurements of OD/ID roundness and tolerances increases with tool wear. Measured roundness and tolerance values increase with tool wear.
4. For economically feasible and statistically reliable measurements using coordinate measuring machines, 12 and 16 sample points are recommended for OD/ID roundness, 3 to 15 sample points for OD/ID tolerances, and 18 to 30 for concentricity of machined cylindrical parts.

ACKNOWLEDGMENTS

The work presented in this paper has been partially supported by the Intelligent Systems Center at University of Missouri-Rolla.

REFERENCES

1. Tipnis, V.A., Report Committee Chairman, "Research Needs and Technological Opportunities in Mechanical Tolerancing-Results of an International Workshop", ASME CRTD-15, ISBN No. 0-7918-0471-2, (1959).
2. Walker, R., "GIDEP Alert from Westinghouse", XI-A-88-01, (1988).
3. Schneiber, Rita R., "The Methods Divergence", Manufacturing Engineering, May (1990).
4. Requicha, A.A.G., "Towards a theory of Geometric Tolerances", International Journal of Robotics Research, Vol.2, No.4, pp. 45-60, (1983).

5. Requicha, A.A.G. and Chan, S.C., "Representation of Geometric Features, Tolerances and Attributes in Solid Modelers Based on Constructive Geometry", IEEE Journal of Robotics and Automation, Vol. Rt-2, No.3, pp. 156-166, (1986).
6. Srinivasan, V. and Jayaraman, R., "Issues in Conditional Tolerances for CAD Systems", IEEE International Conference on Robotics and Automation, St. Louis, MO, (1985).
7. Hocken, R. H., "Sampling Techniques for Coordinate Measuring Machines", an NSF funded project being conducted at the University of North Carolina at Charlotte, Mechanical Engineering and Engineering Science Department, (1989).
8. Brown, Curtis W., "Dimensional Inspection Techniques for Sample Point Measurement Technology", Journal of Precision Engineering, Vol. 14, No. 2, pp 110-111, (1992)
9. Elshennawy, Ahmad K., Ham, Inyong, Cohen, Paul H., "Evaluating the Performance of Coordinate Measuring Machines", Quality Progress, pp. 59-65, (1988).

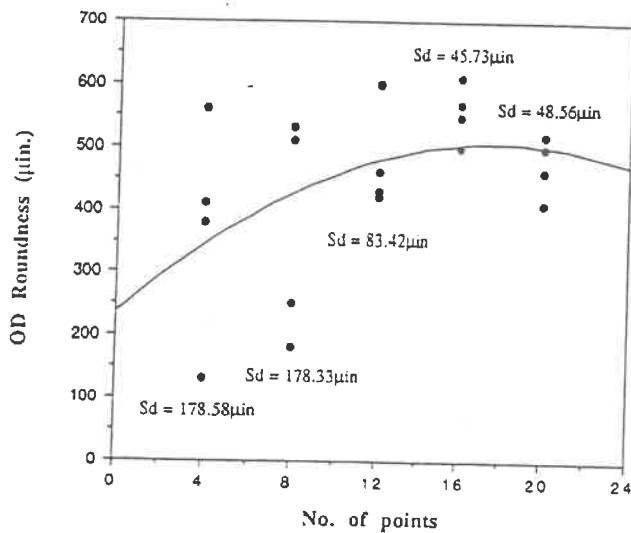


Figure 2: O.D. Roundness vs sample point density

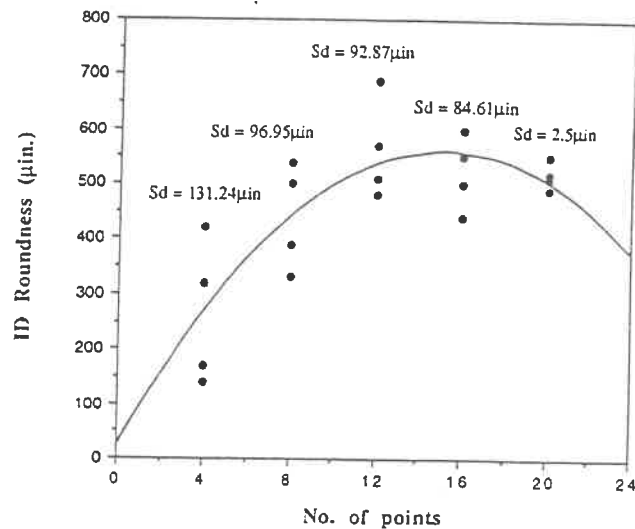


Figure 3: I.D. Roundness vs sample point density

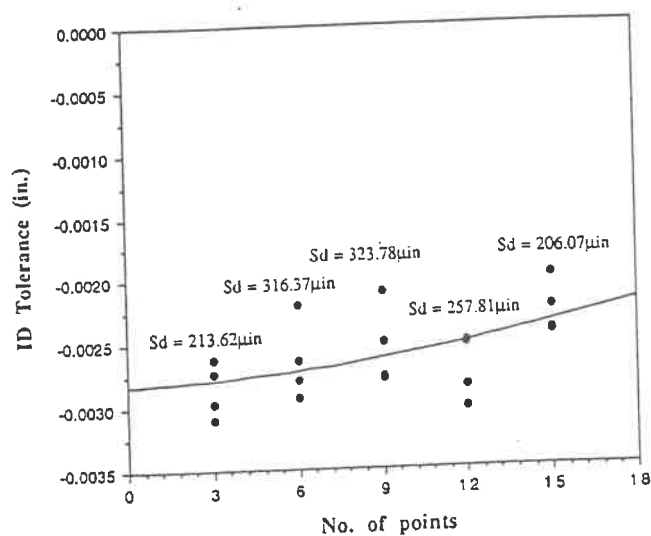


Figure 4: I.D. Tolerance vs sample point density

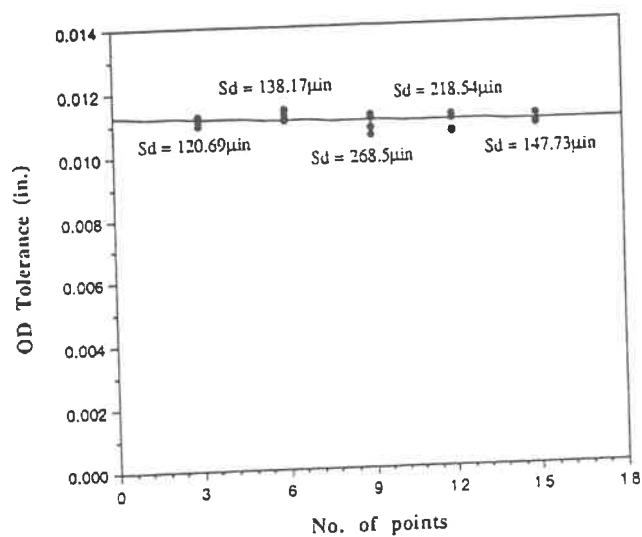


Figure 5: O.D. Tolerance vs sample point density

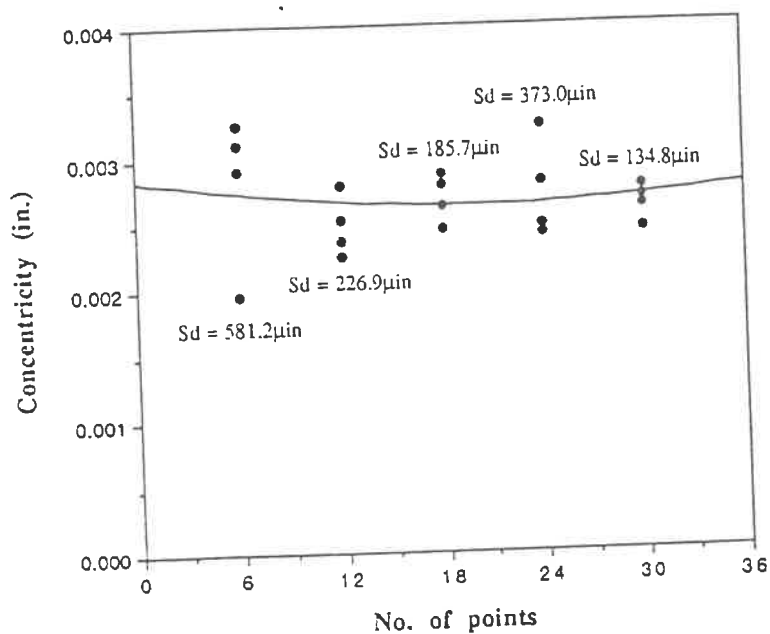


Figure 6: Concentricity vs sample point density

THE EFFECT OF CIRCULAR GROOVE ON FLATNESS ERROR OF METAL MIRROR CAUSED BY CHUCKING FORCE

Norimitsu Ozawa*, Shigeru Kawabe** and Kiyoshi Mizuhara*

*Mechanical Engineering Laboratory, AIST, MITI

1-2 Namiki, Tsukuba, Ibaraki, 305 JAPAN

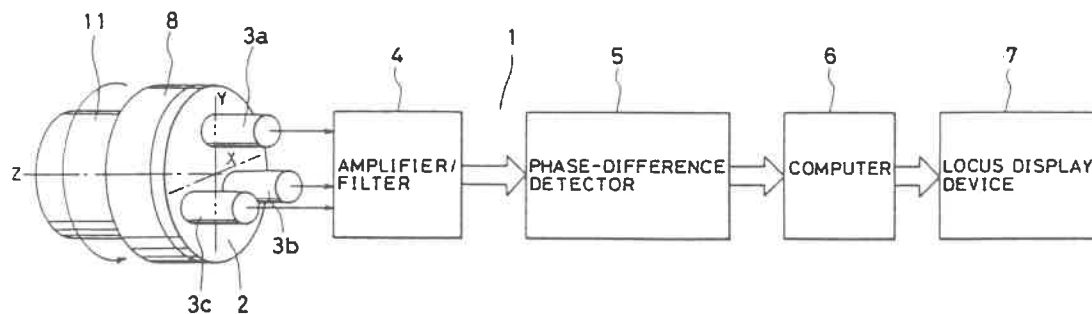
**Precision Technology Research Institute of Nagano Prefecture
9959 Okaya, Nagano, 394 JAPAN

1. Introduction

To produce highly accurate optics in single point diamond turning technology, it is necessary to align between height of the tool edge and axis of spindle rotation. Miller has proposed a tool centering method according to "Chamfered post technique"[1]. For the purpose, however, it is important to develop a detection method of central position of spindle rotation(CPSR). Eiju has developed an optical method to detect the CPSR[2]. We have developed a new measurement method of the CPSR using a precision steel ball and a non-contact displacement sensor[3], and have proposed an automatic detection method using a precision flat mirror and three non-contact sensors[4] illustrated in Figure 1. In order to realize the automatic method, there are two technical problems to be solved. (1)The mirror has to be held at the spindle end without decreasing a flatness accuracy. (2)The mirror must keep flatness under operational conditions of the spindle. Some research concerning these problems has been carried out. Barkman has described a "Workpiece fixturing for precision machining"[5]. Shinno has proposed a distortion-free workpiece fixturing method[6]. This paper describes the effect of clamping force on flatness of the mirror that has a circular groove. Experimental investigations and results of FEM analysis have been shown.

2. Experimental procedure

Three metal mirrors are made of oxygen free copper specimen. Dimension of the mirrors is 30mm diameter and 25mm length. One side surface of every mirror was machined by diamond turning technique and coated with thin gold film after turning. In addition, a 2mm wide circular groove was cut in center position of the mirror body as illustrated in Figure 2. Depth of the groove is 0mm, 2.5mm and 7.5mm respectively. Special three jaw chuck was used in this experiment. To measure the clamping force, small force sensor using a strain gauge was attached to the jaws respectively. Figure 3 shows a part of the chuck and the mirror. As shown in this figure, contact surface between jaw and mirror has the same radius of curvatures and contact area is 54mm² respectively.



(11: Spindle, 8: Stage, 2: Metal mirror, 3: Sensor)

Figure 1. Automatic detection method of the CPSR

Deformation of the mirror surface depends on clamping force and was measured by using a laser interferometer(Zygo PTI). The laser interferometer was mounted on vibration isolator and covered by vinyl sheet to prevent air disturbance on the laser pass. Experimental apparatus was installed in the temperature controlled room.

3. Experimental results and discussion

Changes of the clamping force were observed for 60 minutes since the beginning of the experiment under 500N applied to each jaw. As a result, there was about 1% reduction of applied force. In flat machining by turning operations, generally, it is difficult to make completely flat surface. The cross sectional profiles of mirrors used in this experiment are concave. Furthermore, several burrs or dull part exists the edge of the mirror. Because of these reasons, peripheral part of the mirror is masked on the flatness measurement. Thus, the measuring area became about 64% of the mirror surface. Figure 4 shows relationship between clamping force and flatness of the mirror. As clamping force increase, the flatness is changed for all mirrors. The reason why flatness is changed, central part of concave is swelled by the clamping force. In addition, the resulting peak-to-valley flatness error depends on the groove depth. By comparing two mirrors of groove depth 0mm and 7.5mm, total changes of the flatness of the 7.5mm depth mirror decrease about 1/3 than of the 0mm depth mirror. The larger the groove depth, the smaller influence of the clamping force, but the lower the mirror body stiffness.

4. FEM analysis and discussion

To investigate the deformation behavior of whole surface of the mirror, FEM analysis was carried out according to the conditions shown in Table 1. A typical result of the analysis illustrated in Figure 5. From this figure, in case of the grooved mirror, it has been revealed that deformation of surface of the mirror body may be stopped by the groove. However deformation of contact surface with the jaw of grooved mirror is larger than non-grooved mirror. In addition, deformation of the mirror surface shows different characteristics between two mirrors. In case of the non-grooved mirror, the closer to center of the mirror, the larger deformation. On the other hand, surface of the grooved mirror deformed uniformly. Table 2 shows comparison of experimental results and FEM analysis. Results of the FEM analysis are in good agreement with experimental results.

5. Conclusions

In this paper, the effect of circular groove on flatness error of the metal mirror surface caused by clamping force were discussed. The results can be concluded as follows;

- (1)Influence of the clamping force on the flatness of the mirror can be reduced by the groove.
- (2)Deep groove is more effective, but reduction of stiffness of the mirror body should be considered.

References

- [1] A.C.Miller et.al., Nanometer tool centering for diamond turning applications, Proc. of the ASPE 1992 Annual Meeting, 216
- [2] Eiju,T., et. al., Determination of the central position of the rotating object by Laser Doppler velocimetry, Optical Engineering, Vol.30,No.11(1991) 1825
- [3] Ozawa,N. et. al., A new measuring method of central position of spindle rotation(1st report: Steel ball method), Proc. of the ASPE 1992 Annual Meeting, 285
- [4] Ozawa,N., Method and device for measuring the center of rotation of a rotating shaft, U.S.Patent, No.5,390,151, Feb.14,1995
- [5] W.E.Barkman, Workpiece fixturing for precision machining, Precision Engineering, Vol.4, No.2(1982) 101
- [6] Shinno,H., et. al., Workpiece fixturing methods for ultra-precision machining, Trans. of the JSME, Vol.51, No.471(1985) 3125(in Japanese)

Groove depth d (mm)	0	2.5	7.5
Flatness of mirror (nm p-v)	98.8	76.0	70.4
Surface roughness (Rmax)	0.04 μ m		
Material	Oxygen free copper		

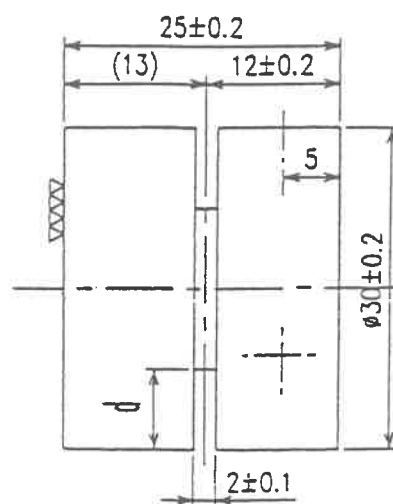


Figure 2. Metal mirrors

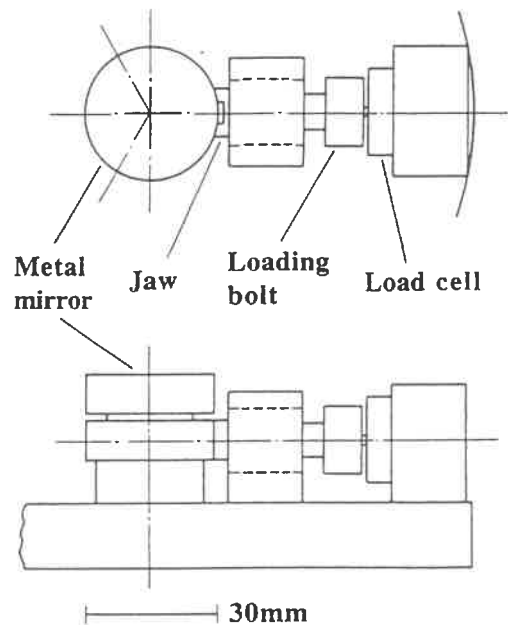


Figure 3. Three jaw chuck

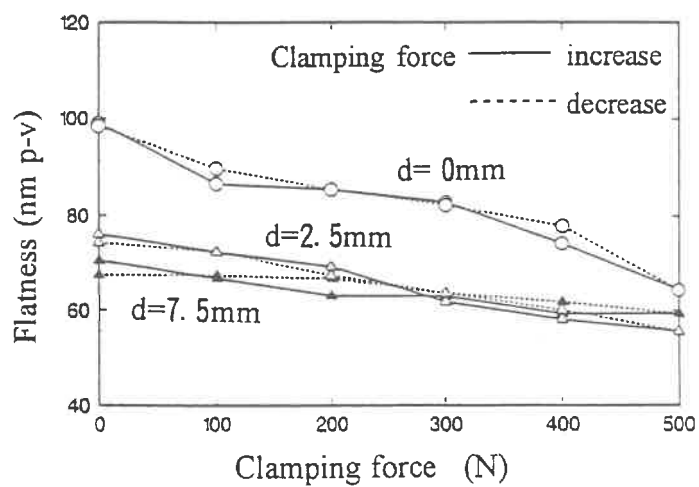


Figure 4. Relationship between clamping force and flatness of the mirror

Groove depth (mm)	0	2.5	7.5	Boundary conditions
Number of node	832	888	888	Constrained nodes: 4 nodes of bottom Loading condition: 500 (N), normal and uniform load
Number of element	567	585	585	

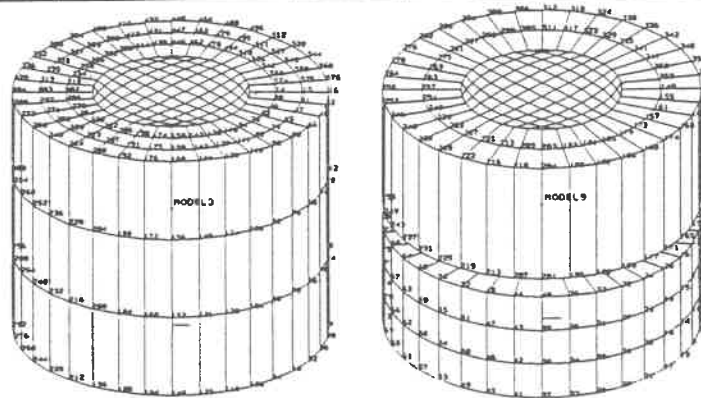


Table 1. Condition of FEM analysis

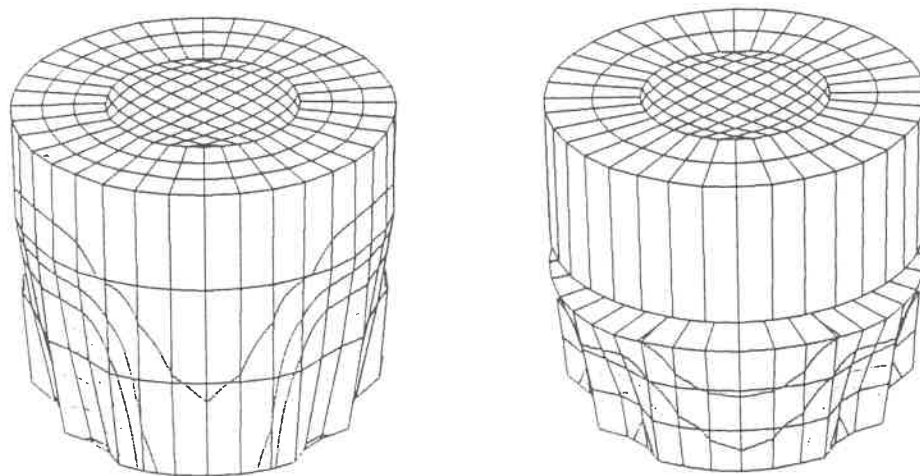


Figure 5. Example of FEM analysis(Left: d=0mm, Right: d=7.5mm)

Groove depth (mm)	0	2.5	7.5
Experimental results (nm)	34.3	19.5	9.6
FEM analysis (nm)	32.9	27.9	8.8

Table 2. Comparison of experimental results and FEM analysis(0-500N)

Sampling Plan for Point-To-Point Performance Test of Touch Trigger Probes on Coordinate Measuring Machines

Yin-Lin Shen and Sungho Moon
Department of Civil, Mechanical, and Environmental Engineering
George Washington University
Washington, DC 20052

INTRODUCTION

Touch trigger probes are devices through which coordinate measuring machines (CMMs) collect dimensional measurement data in modern manufacturing environments. The vast majority of probes used on CMMs are touch trigger probes, and most of these probes employ a kinematic seating mechanism for the probe stylus (Reid, 1993). Point-to-point probing test is typically used to evaluate touch trigger probe performance. The test method specified in the ANSI (American National Standards Institute) B89 standard consists of measuring 49 points on a highly spherical reference ball and the 49 points are distributed by following a sampling plan or test pattern (ANSI, 1991). The user is allowed to specify desired sampling plan which contains 49 points. Standards of other countries usually require that a number of points are taken over one hemisphere of a reference ball roughly uniformly distributed. (For example, fifty points are used in VDI/VDE 2617 standard.) A best-fit sphere is then computed using the manufacturer's recommended algorithms (usually least squares) and the set of radii to each probing point (from the best-fit center) is calculated. The range of the 49 radii is defined as the point-to-point performance index.

In this research, we investigated the 49-point sampling plan specified in the ANSI B89 standard when touch trigger probes are used on CMMs. The sampling plan was tested by using experimental data of a touch trigger probe on a computer-controlled CMM and research findings indicate that the probe performance is overestimated due to a systematic bias in the Z coordinate of the best-fit reference ball center in the probing test. In this paper we propose a two-latitude sampling plan synthesis method based on a pretravel model for touch trigger probes. It is shown that the proposed method can be used to synthesize sampling plans resulting in a near-unbiased reference ball center.

ANSI B89 49-POINT SAMPLING PLAN IN POINT-TO-POINT PROBING TEST

The B89 probe performance test method consists of measuring 49 points on a high-quality reference sphere by following a sampling plan. The user can specify desired sampling plan which contains 49 points. In the B89 default sampling plan, 49 points are probed on the reference ball (test ball) at five different heights on the ball: 12 equally spaced points per latitude are taken on four specified latitudes (approximately 100°, 90°, 60°, and 30°) and one point on the north pole. In addition, there is a 10° rotation of the 12 points on one latitude after the 12 points on the previous latitude are taken. The B89 default sampling plan spreads 49 measurements over the reference ball. It is designed to test three-dimensional (3D) probe performances associated with different probe approach directions.

EVALUATION OF B89 49-POINT SAMPLING PLANS

Experimental probing data were collected on a computer-controlled CMM, and a commonly-used 5-way 3D touch trigger probe with a 50 mm straight stylus extension was used. The probe and the stylus were oriented vertically. In each experiment, 1561 points of probe data, which represent 1561 probe approach directions, were measured. These points are distributed in thirteen (13) latitudes of the reference ball and one point at the north pole. The polar angles of the thirteen latitudes are: 130°, 120°, 110°, ..., 20°, and 10°. One hundred and twenty (120) equally spaced points are measured in each latitude. The point index of the 1561 probe data is arranged as follows: the first 120 points (point 1 to 120) are on the 130° polar latitude, the second 120 points (point 121 to 240) on the 120° polar latitude, ..., and the last point is on the north pole (point 1561). A probe approach speed of 7.5 mm/sec was used. The experiment was repeated ten (10) times. From the experimental data, we extract 100 sets of 49-point data by following the B89 sampling plan: 10 sets of different 49-point data can be extracted from each 1561-point

experiment; and 10 repetitions of the 1561-points experiment provide 100 sets of 49-points data. For each 49-point data set we calculate the best-fit center and radius of the reference ball by using the 49-point data and a least-squares sphere fitting software developed by the National Institute of Standards and Technology (NIST) (Hopp, 1993). We then calculate 49 radii from the measurement points to the best-fit center. We can find the range of the 49 radii, which is the probe performance index, for each of the 100 49-point data sets.

To ensure that we obtain a good estimate of the true center of the reference ball from the 49-point data, we use a dense-data approach to calculate the reference ball center. The dense data are defined as the probing data around the equator of reference ball in the 1561-point experiment, i.e., all data from the 130° latitude up to the 50° latitude (9 latitudes, 1080 points). For each 1561-point experiment, we can calculate a best-fit reference ball center (1080-point center) by using the NIST sphere fitting software. We compare it with the centers of the 10 49-point data sets extract from this experiment. We then follow the same procedures to calculate 49 radii for each 49-point data by using the 1080-point center. The values of the 1080-point reference ball radius are extremely stable (between 7.9881 mm and 7.9883 mm), while those of the 49-point cases fluctuate in a small range of approximately 1 μm (7.9868 mm to 7.9878 mm). As for the best-fit reference ball centers, the X and Y coordinates of the 49-point center are very close to those of the 1080-point center; however, there is a bias in the Z coordinates between the best-fit reference ball centers obtained by using the two methods (see Figure 1).

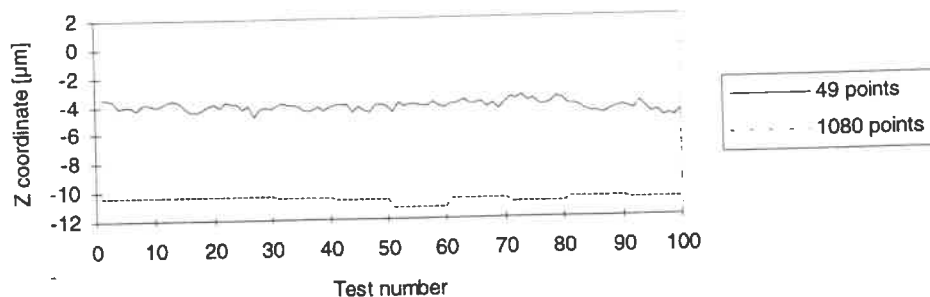


Figure 1. *Z coordinate of the reference ball center using the B89 49 point method and the 1080 point method*

The impact of the quality of the best-fit reference ball center on probe performance is then evaluated. Figures 2 and 3 show the range of the 49 radii (probe performance index) of the 100 test cases by using (1) the 49-point best-fit center and (2) the 1080-point best-fit center. When the 49-point centers are used, the probe performance index (the range of 49 radii) of the 100 test cases has a mean of 10.847 μm and a standard deviation of 1.106 μm . When the 1080-point center are used, the probe performance index (the range of 49 radii) of the 100 test cases has a mean of 16.775 μm and a standard deviation of 1.001 μm . The probe performance is overestimated when B89 49-point sampling plan is used.

TWO-LATITUDE SAMPLING PLAN SYNTHESIS METHOD FOR PROBING TEST

We use a pretravel model for touch trigger probes to develop a sampling plan synthesis method for point-to-point probing tests. Pretravel is the travel distance of the CMM between the instant the probe touches the workpiece and the instant the probe generates a trigger signal. It is highly repeatable but can be adversely influenced by force variations encountered in different probe approach directions in dimensional measurements. Pretravel, mainly caused by bending deflection of the probe stylus shaft, accounts for the majority of touch trigger probes (Jarman and Traylor, 1990; Reid, 1993). Shen and Zhang have developed a pretravel model for vertically-oriented touch trigger probes with straight styli (1996), and the model accounts for bending deflection of the stylus shaft at the trigger instant. The pretravel behavior above and below the equator of the reference ball is approximately symmetrical according to the pretravel model. Based on this near-symmetry characteristic of probe pretravel, we propose a two-latitude sampling plan synthesis method for point-to-point probing test by using the pretravel model.

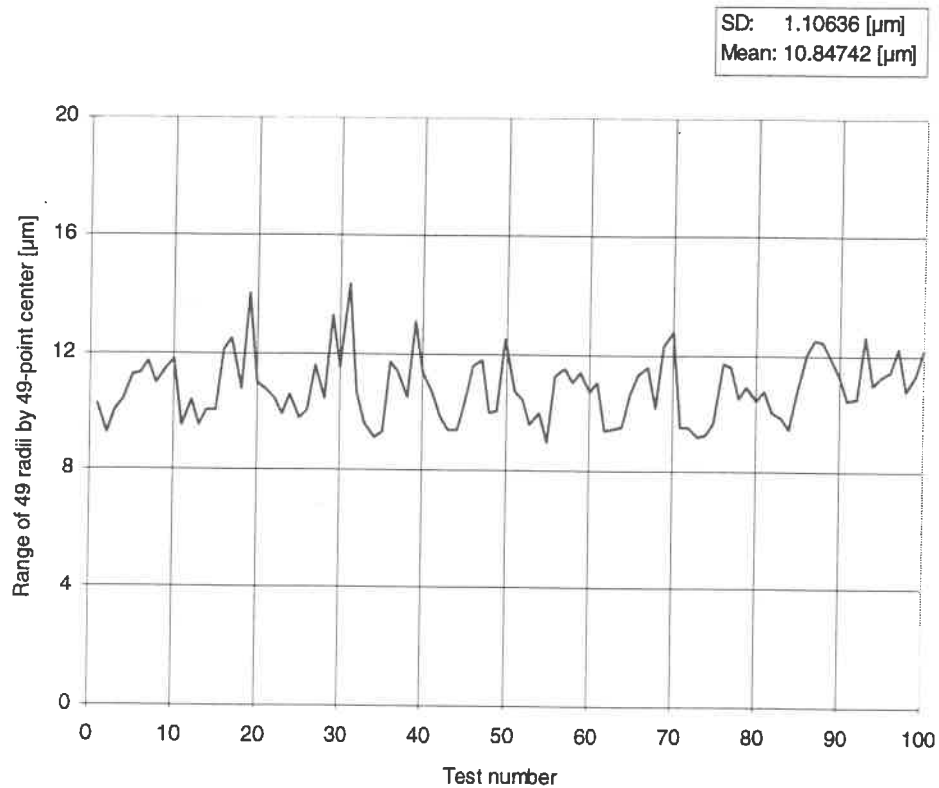


Figure 2. Range of the 49 radii obtained by using the B89 49-point center

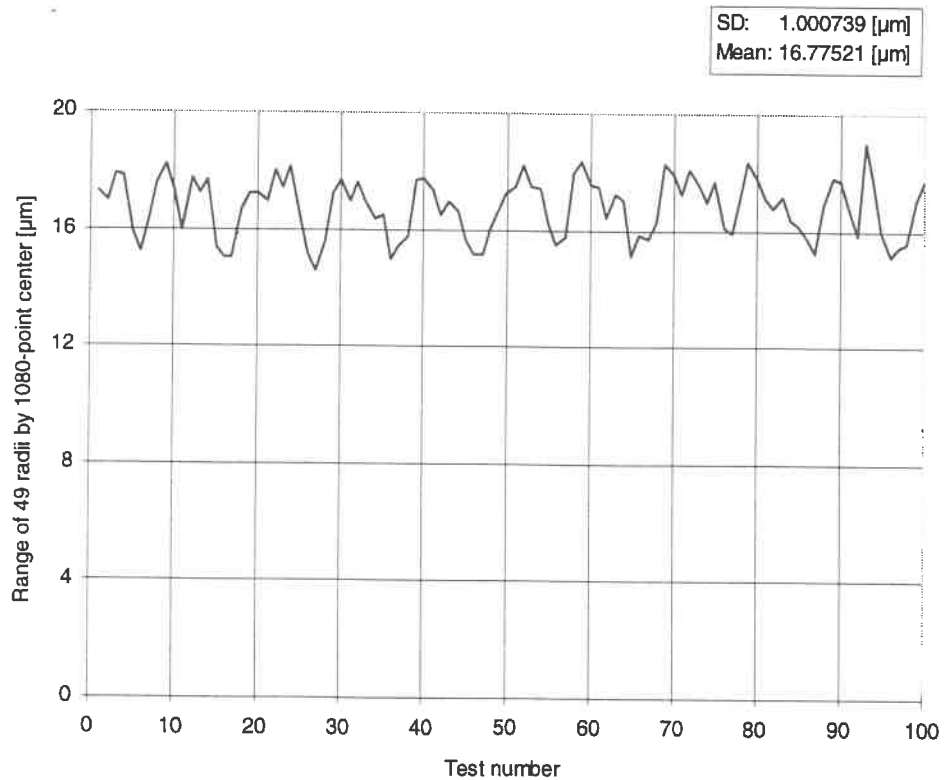


Figure 3. Range of the 49 radii obtained by using the 1080-point center

The two latitudes include one latitude above the equator and another latitude below the equator of the reference ball. We choose one latitude first (e.g., 100° latitude), then use the pretravel model to identify the corresponding latitude which would result in an unbiased reference center. We choose 24 equally spaced points on each latitude, so each sampling plan consists of 48 points on two latitudes. There is no rotation of touch trigger probes, we use 15 different starting angles to generate 15 24-point data sets on each latitude. As a result, we obtain 15 different 48-point data sets for a specific two-latitude sampling plan. We then calculate 15 best-fit centers and radii of the reference ball for each sampling plan. From the dispersion of the 15 best-fit centers, we evaluate the quality of the best-fit reference ball center. We choose 100° latitude first because it is the lowest latitude in the B89 method.

TESTS OF TWO-LATITUDE SAMPLING PLANS WITHOUT NOISE IN THE DATA

Figure 4 shows the least squares fit center (mean and standard deviation) when the 100° latitude is fixed in implementing the two-latitude sampling plan synthesis method. The 86°-100° pair gives a near-unbiased Z coordinate; the mean is 0.004 μm , the standard deviation (s.d.) is 0.020 μm . The X and Y coordinates of the center are very close to the true center; mean is 0.0004 μm and s.d. is 0.0006 μm for X; and mean is 0.0001 μm and s.d. is 0.0005 μm for Y. The calculated radius of the reference ball has a mean value of 7.987924 mm, and the s.d. is 0.068 μm .

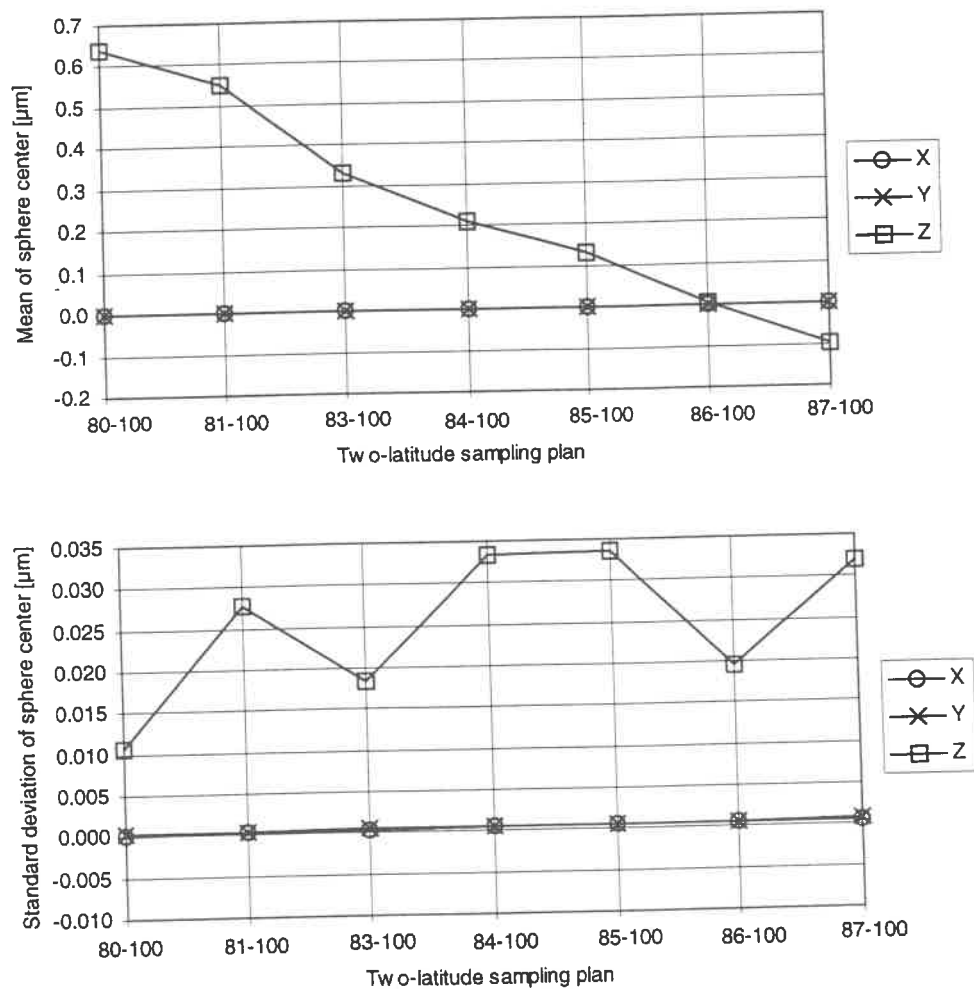


Figure 4. Center of the reference ball by using the 80°-based two-latitude 48-point sampling plan: (a) mean and (b) standard deviation of X, Y, and Z coordinates

The center of the calculated reference ball is very close to the true center location (0,0,0) when the 86°-100° two-latitude sampling plan is used in the probing test. There is no noise in the synthesized measurement data. Tables 1 and 2 summarize the results of other tests: when 110°, 120°, and 130° latitudes are selected in implementing the two-latitude sampling plan; 76°-110° pair, 66°-120° pair, and 56°-130° pair give near-unbiased reference ball center.

Table 1. Mean of the center and radius of the reference ball of the tested two-latitude sampling plans

Two-latitude sampling plan	X [μm]	Y [μm]	Z [μm]	R [mm]
56°-130°	0.000004	0.000005	0.006536	7.990268
66°-120°	0.000007	-0.000002	0.003930	7.989152
76°-110°	0.000065	-0.000020	0.000182	7.988362
80°-106°	-0.000197	0.000026	0.000399	7.988144
86°-100°	0.000397	0.000109	0.004260	7.987924

Table 2. Standard deviation of the center and radius of the reference ball of the tested two-latitude sampling plans

Two-latitude sampling plan	X [μm]	Y [μm]	Z [μm]	R [mm]
56°-130°	0.000036	0.000031	0.006417	0.000055
66°-120°	0.000057	0.000072	0.006260	0.000062
76°-110°	0.000292	0.000268	0.011197	0.000066
80°-106°	0.000729	0.000844	0.009250	0.000067
86°-100°	0.000578	0.000464	0.019569	0.000068

TESTS OF TWO-LATITUDE SAMPLING PLANS WITH NOISE IN THE DATA

To test the robustness of the proposed sampling plan synthesis method, we add Gaussian noise (normally distributed noise) to the 48-point data synthesized by using the pretravel model, and generate 1500 sets of 48-point data containing noise. We use noise data from a normal distribution with mean zero and s.d. of 1 μm in the data generating process; i.e., 99% of the noise falls within $\pm 3 \mu\text{m}$. The 86°-100° pair gives a mean Z coordinate close to zero (mean is -0.016 μm), however, the s.d. is 1.178 μm (as compared to 0.020 μm in the no-noise case). Furthermore, the Z coordinate oscillates between -3.981 μm and +3.770 μm from the simulation results. The mean values of the X and Y coordinates of the center are still very close to the true center; 0.0049 μm for X and 0.0113 μm for Y; however, the s.d. values increase: 0.201 μm for X, and 0.211 μm for Y. The calculated radius of the reference ball still very stable: mean is 7.987920 mm, and s.d. is 0.170 μm .

DISCUSSION AND CONCLUSION

The center of the calculated reference ball center is a critical parameter in the evaluation of probe performance. It is shown that the ANSI B89 49-point sampling plan results in a biased Z coordinate when the 49 points are used to find the best-fit center and radius of the reference ball. The same problem exists in standards of other countries because the specified sampling plans cover only one hemisphere of the reference ball. This finding suggests that the probe performance (the range of the 49 radii) can be overestimated significantly due to the shift of the reference ball in the Z axis.

A two-latitude sampling plan synthesis method, based on a pretravel model for kinematic touch trigger probes, is proposed. The method can be implemented to synthesize two-latitude sampling plans which result in a near-unbiased reference ball center in the probing test. We present the statistics of the calculated reference center to support the suggestion. Furthermore, we use simulations to check the effect of noise to the proposed method by adding Gaussian noise (mean = 0, and s.d. = 1 μm) to each sampled point of the 48-point data. The noise-embedded data reflect real-world situations in which touch trigger probes are used, and the research finding suggests that the uncertainty in the reference ball center contributes to the overall uncertainty of the measurement process. The

uncertainty in identifying the reference ball center should be included in the assessment of the total uncertainty of measurement processes. Moreover, the center shift range increases significantly when noise is added to the synthesized probe data. As a result, we suggest repeated measurements and averaging be used to reduce the range of reference ball center shift in probe performance test and probe calibration, particularly when precision measurement results are required.

The proposed two-latitude sampling plan synthesis method can be used to assess the center of the reference ball in probing tests. However, the method is not designed to replace the 49-point method specified in ANSI B89 standard in that the 49-point test is needed to test probe performances in various approach directions for touch trigger probes in three dimensional (3D) probe tests. In addition, more points (than 48 points used in the tests) and repeated measurements may be useful in ensuring the quality of the probe performance tests.

ACKNOWLEDGEMENT

This work was supported in part by the NIST Computational Metrology Program under Grant No. 70NAB4H1681 and the National Science Foundation under Grant No. DMII-9308648. We thank the Computational Metrology Group of the National Institute of Standards and Technology (NIST) for useful discussions. We also thank Bruce Borchardt of NIST for providing experimental data, and Kusol Prommul, Xianping Zhang, and Marlow E. Springer of the George Washington University for useful discussions.

REFERENCES

- ANSI/ASME B89.1.12M-1990, Methods for Performance Evaluation of Coordinate Measuring Machines, ASME, New York (1990).
- Jarman, T., and Traylor, A., Performance Characteristics of Touch Trigger Probes, *Proceedings of Precision Metrology with Coordinate Measurement Systems* (SME, Troy, Michigan, 1990).
- Reid, C., Probe Technology - Beyond Accuracy, *Proceedings of Applying Imaging and Sensing Technology to CMM Applications* (SME, Nashville, Tennessee, 1993).
- Bosch, J., Evolution of Measurement, in *Coordinate Measuring Machine and Systems*, J. Bosch, Ed., Marcel Dekker Inc., New York (1995).
- Butler, C., An Investigation Into the Performance of Probes on Coordinate Measuring Machines, *Industrial Metrology* (v2, 1991) pp. 59-70.
- Hopp, T., Computational Metrology, *Proceedings of the International Forum on Dimensional Tolerancing and Metrology*, pp. 97-111 (1993).
- Shen, Y., Zhang, X., A Pretravel Model for Touch Trigger Probes with Straight Styli, Part I: Vertical Probes, *Proceedings of the North American Manufacturing Research Conference*, pp. 127-132, Ann Arbor, MI, May, 1996.
- Estler, W.T., Phillips, S.D., Borchardt, B., Hopp, T., Witzgall, C., Levenson, M., Eberhardt, K., McClain, M., Shen, Y., and Zhang, X., Error Compensation for CMM Touch Trigger Probes, *Proceedings of ASPE Annual Meeting* (Austin, Texas, 1995), pp. 428-431.

A Robust Pretravel Model for Touch Trigger Probes in Coordinate Metrology

Yin-Lin Shen and Marlow E. Springer

Department of Civil, Mechanical, and Environmental Engineering
George Washington University
Washington, DC 20052

INTRODUCTION

This paper describes a pretravel model for kinematic touch trigger probes widely used on coordinate measuring machines (CMMs). A trigger signal is generated when the probe stylus touches the workpiece in the measurement process and the signal enables the recording of an x-y-z coordinate point on the workpiece surface where the probe stylus touches. In fact, the trigger signal is not generated at the touch instant, and the probe continues to travel in the probe approach direction and the stylus continues to be deflected until a probing force (on the probe stylus) is large enough to set off the trigger signal at the trigger instant. The travel distance between the touch instant and the trigger instant is known as probe pretravel. Pretravel is a well-known type of error, and accounts for the majority of touch trigger probe errors (Jarman and Traylor, 1990; Reid, 1993). It is a large source of measurement uncertainty in touch-trigger-probe CMMs (Butler, 1991; Bosch, 1995; NIST, 1996). Pretravel errors are highly repeatable but can be adversely influenced by force variations encountered in different probe approach directions (Reid, 1993). This can lead to measurement errors caused by variations in stylus shaft bending prior to triggering. Pretravel errors are referred to as probe-lobing errors because the probe yields a three-lobed error pattern when inspecting a ring gauge.

Estler et al. (1995 and 1996) have developed a multi-parameter pretravel model accounting for the combined effects of rotational displacement, bending displacement, and frictional effects. Shen and Zhang (1996a and 1996b) have developed pretravel models accounting for bending of the stylus shaft for vertically- and horizontally-oriented touch trigger probes. In this paper, a trigger force model for vertical probes was first derived and used to model the trigger force, i.e., the probing force at the trigger instant. A basis for this work was the assumption that the probe behaves statically as a cantilever beam. The trigger force model was used to model the bending deflection of the stylus shaft at the trigger instant in order to verify the cantilever beam assumption. Experimental data of trigger forces and x-y-z coordinates were used to validate the pretravel model. It is shown that 24 points of equatorial data measured in the probe calibration process can be used to accurately identify the only model parameter. Also, 12 points of equatorial data provides reasonably good results.

PRETRAVEL MODEL

When the probe tip is in contact with the workpiece during the measurement process, a probing force is developed. This probing force causes the stylus shaft to deflect, resulting in the probe pretravel. Calculation of the deflection is dependent upon the magnitude and direction of the trigger force. The trigger force is defined as the probing force at the trigger instant. The free body diagram of the probe tripod-stylus is shown in Figure 1. The spring force, F_s , and the weight of the tripod-stylus structure, W , act downward (negative Z direction) to keep the tripod in its rest position. Each tripod leg (A,B,C) of length R is supported by two of six contact forces (F_{A1} , F_{A2} , F_{B1} , F_{B2} , F_{C1} , F_{C2}) between the tripod and the kinematic seating arrangement. A kinematic support angle, α , can be used to describe the direction of the contact forces. The probing force, F_p , increases from zero before the probe tip contacts the workpiece, to a maximum value at the trigger instant. The direction of F_p is opposite to the probe approach direction, described by the polar angle (θ) measured from the positive z-axis to the probing force, and the azimuthal angle (ϕ) measured from the positive x-axis to the component of the probing force in the x-y plane (positive from the positive x-axis to the positive y-axis).

During the duration of the probing, from the contact instant to the trigger instant, the following forces are involved: F_s , W , F_{A1} , F_{A2} , F_{B1} , F_{B2} , F_{C1} , F_{C2} and F_p . Based on kinematic principles six equilibrium equations can be used to describe the relationships among these forces until the trigger instant. Manipulation of these equations leads to solutions for the six contact forces as functions of spring force F_s , weight of the tripod-stylus structure W , kinematic support angle α , probing force F_p , stylus length L , tripod leg length R , polar angle θ , azimuthal angle ϕ , and tripod phase angle ϕ_0 (Shen and Zhang, 1996a). As F_p increases during the measurement process, the six contact forces increase or decrease in magnitude, but remain constant in direction due to the

The diagram illustrates the force system on a ship. A vertical Z-axis represents the ship's longitudinal axis, with the origin at the center of gravity. The X and Y axes are horizontal. A dashed ellipse represents the ship's hull cross-section. Forces are shown as vectors originating from the center of gravity: F_{s+W} (weight and buoyancy) acts vertically upwards; R (resistance) acts horizontally along the negative X-axis; F_{A1} and F_{A2} (air forces) act on the upper hull; F_{B1} and F_{B2} (buoyancy forces) act on the lower hull; F_{C1} and F_{C2} (center of buoyancy forces) act on the center of buoyancy. The angle ϕ_0 is the heel angle. The angle θ is the pitch angle, and ϕ is the roll angle. The 'approach direction' is indicated at the bottom.

At the junctions between two regions, two contact forces will reach the threshold. The probe trigger condition is assumed as follows: when one of the six contact forces reaches F_t , the trigger signal is generated. The trigger force can be expressed as:

$$F_p = \frac{6\sin\alpha F_t - (F_s + W)}{L\sin\theta[\cos(\phi - \phi_0) + \sqrt{3}\sin(\phi - \phi_0)] - \cos\theta - \sin\theta\tan\alpha[\sin(\phi - \phi_0) - \sqrt{3}\cos(\phi - \phi_0)]} \quad (180^\circ \leq \phi - \phi_0 \leq 240^\circ), \quad (4)$$

$$F_p = \frac{6\sin\alpha F_t (F_s + W)}{R \sin\theta [\cos(\phi - \phi_0) + \sqrt{3}\sin(\phi - \phi_0)] - \cos\theta + \sin\theta \tan\alpha [\sin(\phi - \phi_0) - \sqrt{3}\cos(\phi - \phi_0)]} \quad (240^\circ \leq \phi - \phi_0 \leq 300^\circ), \quad (5)$$

$$F_p = \frac{F_s + W - 6\sin\alpha F_t}{R [\cos\theta + 2\sin\theta \cos(\phi - \phi_0)] - 2\sin\theta \sin(\phi - \phi_0) \tan\alpha} \quad (300^\circ \leq \phi - \phi_0 \leq 360^\circ). \quad (6)$$

The deflection of the stylus shaft at the probe tip is caused by bending and compression or tension due to the increasing probing force. The deflection model can be simplified by treating only the dominant bending deflection caused by the horizontal component of the trigger force (Jarman and Traylor, 1991; Shen and Zhang, 1996a) specified by

$$F_H = F_p \sin\theta, \quad (7)$$

where F_H is the horizontal trigger force component. Thus, treating the stylus as a cantilevered beam, the magnitude of the projection of the 3D pretravel vector along the probing force direction is described as

$$\delta_{pt} = \frac{F_H L^3}{3EI} \sin\theta = \frac{F_p (\sin\theta)^2 L^3}{3EI}, \quad (8)$$

where δ_{pt} is the model predicted pretravel, E is the modulus of elasticity of the stylus shaft material, and I is the moment of inertia of the stylus shaft.

EXPERIMENTAL DATA

Experimental probing data were used to determine the magnitude of the threshold contact force (F_t) needed to generate the trigger signal. Using the model of the trigger force (Eqs. 1 to 6), various values of F_t were tested in an attempt to match the shape and magnitudes of the experimental probing force data. Using a value of 3.8 grams, the predicted forces are fairly close to the data, so it is concluded that the threshold force is somewhere close to 3.8 grams. This value yielded acceptable results in the model.

Data were collected using a direct computer controlled (DCC) CMM, and a commonly-used 5 way 3D probe with a 50 mm straight stylus extension. The stylus material is steel and the 50 mm stylus extension has the following geometric configuration: 3 mm diameter (20 mm length), tapered to 2 mm diameter (10 mm length), and 2 mm diameter (20 mm length). The probe and stylus were oriented vertically. The probe approach speed was 7.5 mm per second. In each experiment 1561 coordinate points representing 1561 probe approach directions to a reference sphere were measured. The points were obtained using the following scheme: one hundred twenty (120) points uniformly spaced at each of thirteen polar latitudes of the reference ball (130° , 120° , 110° , ..., 20° , and 10°), plus one point at the north pole. The points were taken beginning at the 130° degree polar latitude and ending at the north pole. The 1561 point experiment was repeated ten (10) times, so we have 10 single runs of 1561 point experimental data. The experimental pretravel data (1561 points) of each single run were calculated using the x-y-z coordinate probing data. Define r_i , $i=1, 2, \dots, 1561$, as the distance from each measured point to the reference ball center. Let R_0 (7.9993 mm) be the sum of radii of the reference ball (4.9996 mm) and the probe tip (2.9997 mm). Pretravel associated with each probe approach direction is then the difference of the measured point distance (r_i) and the reference radius (R_0).

$$\delta_i = R_0 - r_i, \quad (9)$$

where δ_i is the experimental pretravel.

RESULTS

The moment of inertia (I) is needed to predict the pretravel deflections (see Equation 8). Due to the structure of the stylus shaft (varying cross-section, threads, through holes) it is not feasible simply to calculate the moment of inertia. Thus, it is the only unknown parameter of the pretravel model, and the effective radius of the stylus shaft

(r_e) is used to represent that parameter. The least-squares method is used to minimize the error between the experiment pretravel and the model-predicted pretravel. With only one parameter to be determined, there is a closed-form solution.

The baseline model by which to judge the results of the succeeding models was developed using 1080 points of averaged pretravel (from 130 degree latitude to 50 degree latitude, inclusive) to find the model parameter. To reiterate, the other parameters used were: $F_s = 156.6$ gm, $\alpha = 41.0^\circ$, $L = 59.0$ mm, $R = 5.0$ mm, $R_o = 7.9993$ mm, $F_t = 3.8$ gm, and $E = 200.0$ GPa. Using this model, r_e was found to be 1.302 mm. Figure 2 shows the averaged experiment pretravels, the model predicted pretravels, and the prediction errors (model - experiment) for all 1561 approach directions. The mean of the 1561 prediction errors is 0.15 μm , and the standard deviation (std) of the 1561 errors is 0.56 μm .

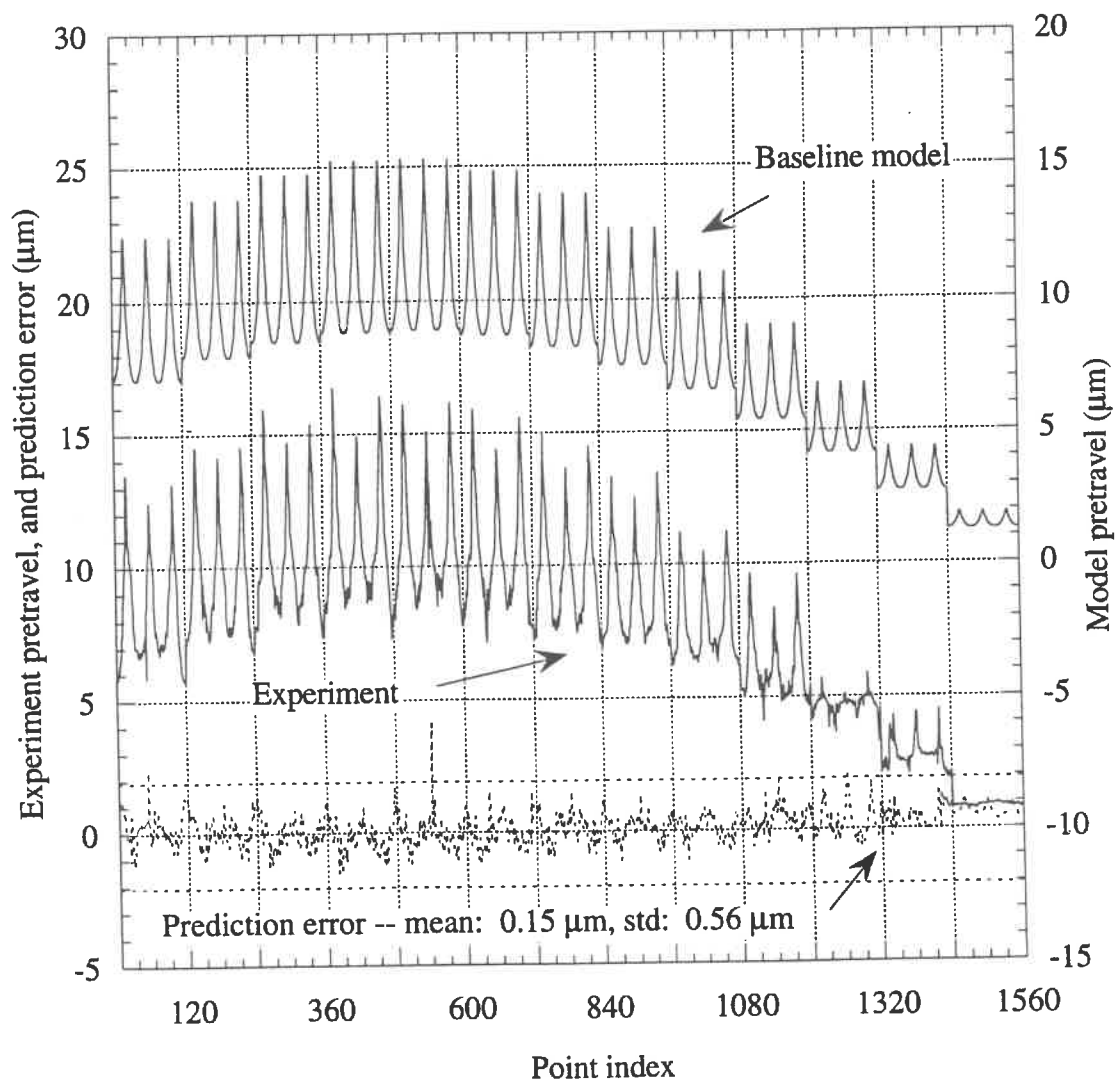


Figure 2. Experiment and baseline model predicted pretravels and prediction error in 1561 approach directions

The model was tested with fewer points, as it is desirable to use the smallest number of points possible to identify the model parameter. The results obtained for these sampling patterns are shown in Table 1. The most

promising testing scheme is that using 24 points equally spaced around the equator. This yields five sample sets from each of the ten 1561 point experimental data sets, making a total of fifty (50) test cases. The model parameter is fairly stable, with a mean of 1.299 mm, and the standard deviation being 0.006 mm. The standard deviations of 1561 prediction errors of the fifty 24 point cases range from 0.52 μm to 0.72 μm . As a check of the robustness of the model, 12 points from the equator were sampled for use in the model parameter calculation. The twelve points were equally spaced, and resulted in ten sample sets from each of the ten data sets, yielding a total of one hundred (100) test cases. There is a gain using these cases, as fewer points must be measured, but there is a cost in the accuracy of the pretravel prediction. The standard deviations of 1561 prediction errors of the 100 point cases range from 0.67 μm to 1.06 μm . The effective radius still has a mean of 1.304 mm, but the standard deviation has increased to 0.009 mm (as compared to 0.006 mm in 24 point cases).

Table 1. *Results of model tests from sparse data cases*

No. of points used in model	Statistic	Phase angle (°)	Effective Radius (mm)	Mean of 1561 errors (μm)	Std of 1561 errors (μm)	No. of runs
24 (equator)	Mean	1.411	1.299	-0.23	0.57	50
	Std	1.165	0.006	0.14	0.03	
	Min	-2.166	1.285	-0.56	0.52	
	Max	3.932	1.315	0.17	0.72	
12 (equator)	Mean	1.324	1.304	-0.10	0.67	100
	Std	3.202	0.009	0.21	0.11	
	Min	-5.358	1.283	-0.63	0.52	
	Max	9.975	1.339	0.68	1.06	

DISCUSSION AND CONCLUSION

By way of this research an analytical model for the probe triggering process has been built, which has deepened the understanding of a key technology in modern dimensional metrology and automation. The threshold contact force level (F_t) is one of the key issues in developing the pretravel model, and experimental probing force data were used to determine the value of this threshold. Use of a correct threshold force level results in more accurate prediction of the probing forces. The model presented herein makes use of data from the equator to determine the model parameter and the phase angle. The beauty of the model is that the probing pretravel behavior is effectively captured by the model without the need to use pretravel data from latitudes other than the equator. Moreover, the model parameter (r_e) is accurately and consistently identified using only twenty-four (24) points from one level. As measurements from only one level are needed, the model can be fit using a 2D artifact, such as a ring gauge, rather than a 3D artifact.

The tripod phase angle must be used in the model, and can be determined using the same equatorial data at no extra cost to the user. However, we have shown that a more accurate value of the phase angle can improve the prediction effectiveness and reduce probing error in CMM measurements. As a result, we suggest that probe vendors provide the phase angle information in ways such as markings on the probe or a signal to the machine system. The model is practical in its simplicity and that it can be implemented by using a computer program, which can be incorporated into the machine's probe calibration process. This provides for real time pretravel compensation for CMMs and machine tools.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation under Grant No. DMII-9308648 and the National Institute of Standards and Technology (NIST) Computational Metrology Program under Grant No. 70NAB4H1681. We thank Professor James D. Lee of the George Washington University (GWU) for useful suggestions, and the Computational Metrology Group of NIST for valuable discussions. We also thank Bruce Borchardt of NIST for providing experimental data, and Xianping Zhang of GWU for discussion.

REFERENCES

- Bosch, J., 1995, *Evolution of Measurement, in Coordinate Measuring Machine and Systems*, J. Bosch, Ed., Marcel Dekker Inc., New York.
- Butler, C., 1991, "An Investigation Into the Performance of Probes on Coordinate Measuring Machines", *Industrial Metrology*, Vol. 2, pp. 59-70.
- Estler, W.T., S.D. Phillips, B. Borchardt, T. Hopp, C. Witzgall, M. Levenson, K. Eberhardt, M. McClain, Y. Shen and X. Zhang, 1995, "Error Compensation for CMM Touch Trigger Probes", *Proceedings of ASPE Annual Meeting*, Austin, Texas, pp. 428-431.
- Estler, W.T., S.D. Phillips, B. Borchardt, T. Hopp, C. Witzgall, M. Levenson, K. Eberhardt, M. McClain, Y. Shen and X. Zhang, 1996, "Error Compensation for CMM Touch Trigger Probes", accepted for publication in *Journal of Precision Engineering*.
- Jarman, T., and Traylor, A., 1990, "Performance Characteristics of Touch Trigger Probes", *Proceedings of Precision Metrology with Coordinate Measurement Systems*, SME, Troy, Michigan.
- NIST, 1996, "Superfit Greatly Reduces Major Source of CMM Errors", Press Release 96-06, National Institute of Standards and Technology (NIST), Gaithersburg, MD.
- Reid, C., 1993, "Probe Technology - Beyond Accuracy", *Proceedings of Applying Imaging and Sensing Technology to CMM Applications*, SME, Nashville, Tennessee.
- Shen, Y., and Zhang, X., 1996a, "A Pretravel Model for Touch Trigger Probes with Straight Styli, Part I: Vertical Probes", *Proceedings, North American Manufacturing Research Conference*, Vol. 24, SME, Ann Arbor, Michigan, pp. 127-132.
- Shen, Y., and Zhang, X., 1996b, "A Pretravel Model for Touch Trigger Probes with Straight Styli, Part II: Horizontal Probes", *Proceedings, North American Manufacturing Research Conference*, Vol. 24, SME, Ann Arbor, Michigan, pp. 133-138.

CALIBRATION METHODS FOR DUAL PROBE CMMs

B. P. Vishwas and R. G. Wilhelm

Precision Engineering Laboratory

Department of Mechanical Engineering and Engineering Science

University of North Carolina at Charlotte NC 28223

A low cost method for 2-D software compensation of a dual probe coordinate measuring machine is described and the results are compared with results from a compensation technique using a laser interferometer. The measuring machine (WERTH Inspector) is equipped with both contact and non contact probes; the contact type being a touch trigger probe (TTP) and the non contact type being a CCD camera fitted with lenses of magnification 0.2x to 100x. To make use of the dual probe nature of the measuring machine, two types of calibrated artifacts are planned for calibration; one being a hole plate with ring gages as the holes and the other being a grid plate. The TTP probe and the CCD camera are used to measure the hole plate and the grid plate respectively in different orientations and the measured data is used for 2-D calibration of the measuring machine by two methods.

Early work in the area of calibration of coordinate measuring machines was done at NIST by Hocken and Borchardt[1]. They presented a simple method for the determination of scale and squareness errors in two dimensional measurements. A method for applying software compensation to a commercial three axis coordinate measuring machine was presented by Zhang and Veale[2]. They presented a technique of compensation for geometric positioning errors and some thermal effects. Also, in the area of vision based systems, calibration methods have been developed by Raugh[3] who developed a method for calibration of an e-beam lithography machine (with a high precision X-Y stage) using a movable grid plate and Ananda[4] who presented a method for 2-D calibration of a vision CMM using a 1-D calibrated grid plate. A method for 3-D calibration of a three axis CMM by measuring 2-D reference objects in four positions has been developed at PTB by Kunzmann[5].

As a first step a mathematical model is developed to describe all the parametric error terms associated with the motion of the machine. The 21 geometric errors associated with the CMM are measured using a laser interferometer. The measured data is analyzed and a polynomial is approximated to each type of error data using the least squares method. Direct error compensation is done by linear interpolation all along the measuring volume. In the second step, a set of holes on a calibrated hole plate are to be measured in four different orientations using the TTP probe. The measurements are then repeated for a series of cross patterns on a grid plate using the CCD probe.

In each orientation, the machine coordinate system must be made to coincide with the artifact coordinate system. Using the data in each orientation, a set of equations is formed to relate positions expressed in the machine coordinate system to positions expressed in the hole plate coordinate system by the transformation matrix containing the coefficients of error polynomials. These coefficients are obtained using least squares elimination by minimizing the chi-square.

The measuring machine (WERTH Inspector) belongs to XYZ type (notation introduced by Charleton) where the XY table positions the part, the TTP and vision probe being mounted on the Z-rail. The error model of the CMM is developed by following the method presented by Zhang[2]. The main assumptions made are that the machine components (base, carriages, etc.) nearly behave as rigid bodies and the error terms of a carriage depend only on the position of that carriage, not on the position of any other carriages. The matrix equation describing the true position of the probe with respect to the part is

$$X\vec{Y}Z = R_x[R_y(R_z^{-1} \cdot \vec{P} + \vec{Z} - \vec{Y}) - \vec{X}]$$

After performing the above operations the equations for the compensated coordinates of the CMM are obtained which is of the form

$$X_{Actual} = X_{CMM} + X_p + \Delta X$$

with similar equations for the Y and Z axes. However the Z axis equation is not considered in the model since 2-D calibration is the criterion.

Measurements are carried out with the probe in the vertical position (probe axis perpendicular to XY plane) and hence the probe offsets X_p & Y_p are equal to zero and only Z_p exists. Therefore the error terms for this calibration include all terms except the errors that are function of Z.

Individual parametric errors are expressed in the form of polynomials, from the knowledge of error data. The error data is analyzed and a polynomial is approximated for each type of error. For example the scale error in X axis is expressed as a first order polynomial in X $\delta x(X) = A + BX$.

Figures 1 through 4 show the configurations of the hole/grid plate. The axes of the machine coordinate system are indicated as (xm,ym) and (xg,yg) represent the axes of the hole/grid plate coordinate system. Positions labeled 1 through 9 are hole/grid points.

As a first step, the machine coordinate system is made coincident with the hole/grid plate coordinate system with the axes of both coordinate system being aligned. The next position is a clockwise rotation of the grid plate through 90 degrees and the third position is the rotation of the grid plate through 180 degrees clockwise from the first position. The fourth position is obtained by translating the grid plate in the X direction by a known distance.

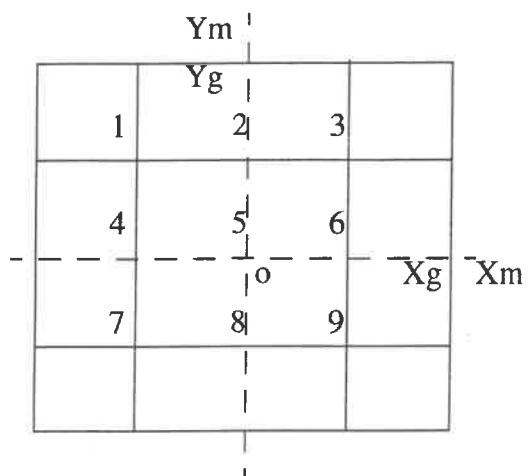


Figure 1. First position.

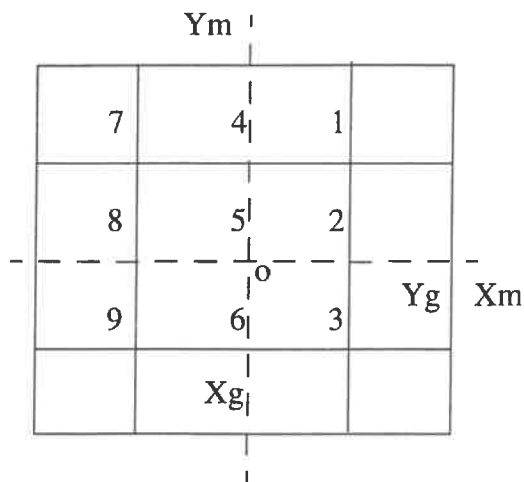


Figure 2. Second position. Rotation in clockwise direction through 90 degrees from the first position.

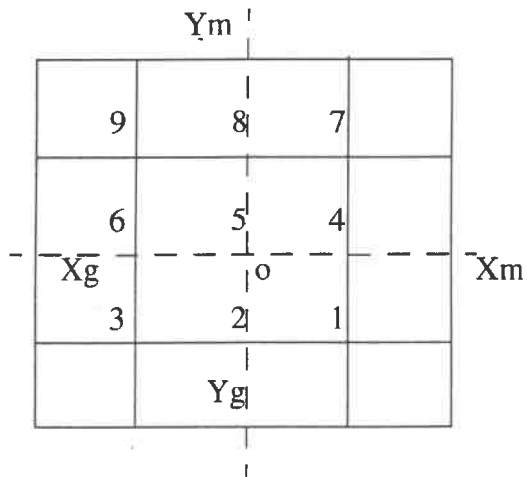


Figure 3. Third position. Rotation in clockwise direction through 180 degrees from the first position.

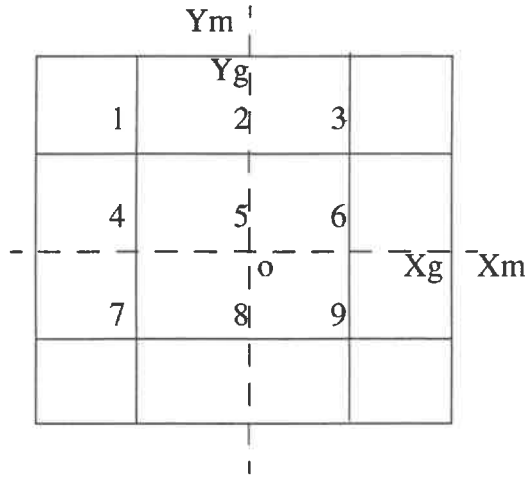


Figure 4. Fourth position. Translation through a known distance along X axis from the first position.

The calibration parameters are developed by relating positions in the hole/grid plate and machine coordinate systems while accounting for the parametric errors of the machine. A typical equation for the relationship between the hole/grid point at position one and position two is shown below.

$$M_{1j} \begin{bmatrix} xm_{1j} \\ xm_{1j} \end{bmatrix} = R_1^{-1} M_{2j} \begin{bmatrix} xm_{2j} \\ ym_{2j} \end{bmatrix}$$

The M matrices refer to the machine errors found at two different positions. The x parameters are the hole/grid points and the R matrix describes the rotations for the various calibration positions. Additional equations are developed by equating the coordinates in the first position to the third and fourth positions respectively. Since the plate is calibrated the distance between any two hole centers is known.

The system of equations is then solved to determine the coefficient of the polynomials by using the least squares method to minimize the chi-square. The main criteria is to minimize the number of features, holes/grid points, and also meet the accuracy requirements.

Experiments are now in progress to determine the minimum feature requirements and to compare the calibration results obtained by hole plate and grid plate with that of software correction by the laser interferometer method. By this method a 2-D calibration using measured artifacts in different orientations will be possible. It is expected that the artifact method will be both faster and less expensive than interferometer method.

Acknowledgments

This work has benefited greatly from the exceptional research environment provided by the UNCC Precision Engineering Laboratory and the Cameron Center for Applied Research. This material is based upon work supported by the National Science Foundation under Grant No. DMI-9457168.

References

1. R. J. Hocken and B. Borchardt, "On Characterizing the Measuring Machine Geometry", NBSIR, 79-1752.
2. G. Zhang and R. Veale, "Error Compensation of Coordinate Measuring Machines", Annals of the CIRP, Vol.34/1, 1985.
3. M. Raugh, "Absolute two-dimensional sub-micron metrology for electron beam lithography", Integrated Circuit Metrology 2, SPIE, Vol 480, 1984.
4. M. V. Ananda, "Residual Error Compensation of A Vision Based Coordinate Measuring Machine", Master's Thesis, UNC Charlotte, 1992.
5. H. Kunzmann, "A uniform concept for Calibration, Acceptance Test, and Periodic Inspection of Coordinate Measuring Machines using Reference Objects", PTB, 1990.

EVALUATION OF ELLIPTICAL PROFILES AND ITS APPLICATION IN CYLINDER MEASUREMENT BY CMMS

Yiding Wang, Kee S. Moon and John W. Sutherland

Department of Mechanical Engineering-Engineering Mechanics
Michigan Technological University
Houghton, MI 49931

ABSTRACT

This paper presents improved algorithms for the evaluation of elliptical profiles measured by coordinate measuring machines. The value D/R , which D is the distance between the center of the ellipse and the origin and R is the half major axis, is identified as a critical value to make the commonly used least square algorithm not robust. The improved algorithm can reduce the computation error caused by singularity of the coefficients matrix when using the least square method. The method can be further applied in the evaluation of the data for cylinder measurement by CMMs when the axis of the cylinder is not parallel to the Z axis of the CMMs. The evaluation errors relating to the distribution of the sampling points and the deviation of the axial direction of the cylinder from the Z direction are investigated by simulation.

Keywords: metrology, CMM, algorithm.

INTRODUCTION

Coordinate measuring machines (CMMs) assess the parameters of mechanical parts on a point-by-point basis. The acquired data are analyzed to produce a "substitute geometry" to be compared with the design intent. However, CMMs based dimensional inspection has some drawbacks. Besides the hardware error of the machine, the most significant ones are the effects of sampling strategies and the fitting algorithms on the reported results. These errors are caused by estimating the measured features according to a particular number and distribution of the measurement points instead of all the points on the surface. Therefore, the reported results may contain large errors. With the development of software technique, the errors caused by the software of CMMs aroused great concerns in coordinate metrology area. As reported by Waldele (Waldele, 1993), certain algorithms are capable of stating that the measurement is worse than the actual data gathered up to an error of 37%. The software errors are usually caused by the fitting objectives of a problem, optimization methods used, and computing environment (word length, rounding method), etc., (Hopp, 1993) to make the evaluation software far from robust.

This paper describes the analysis of the evaluation errors for measuring a cylinder subjected to different sampling conditions in the case that its axis is unparallel to the Z axis of the CMM. The evaluation of elliptical profiles is suggested to be used in the evaluation of a cylinder measurement in this case. The drawbacks of the commonly used least-square fitting algorithms (especially for the singularity problem) are investigated. Improved algorithms for the evaluation of elliptical profiles are presented. The methods can be used in the evaluation of the data acquired by CMM for the workpiece with elliptical profiles.

ANALYSIS OF THE ERRORS IN CYLINDER MEASUREMENT

The most common way of measuring a cylinder is to measure a number of circular traces to be taken at different horizontal sections of the cylinder. For the simplest measurement, three measurement points on each section are needed to fit a circle. The measurement points are usually uniformly distributed. However, in the case that the axis of the cylinder is not parallel to the Z axis, the horizontal sections of the cylinder are not circular but elliptical. Using circles as fitting objectives will cause significant errors. Phillips (Phillips, 1993) illustrated the distribution of distances between the best fit center location and the true center position for a randomly oriented circular feature having a 10 μm elliptical form error. The results show that the correct center is never found using three measurement points (120° spaced) to fit for a circle.

In order to demonstrate the error caused by the effect of the unparallelness between the axis of the cylinder and the Z axis as well as the starting angle of the measurement on the fitting results, a simulation procedure was developed. Three uniformly spaced points in a horizontal section of a cylinder are sampled each time and used to fit for a circle. The angle between the axis of the cylinder and the Z axis is denoted ϕ and the starting angle for each sampling process is denoted θ (Fig. 1). The relative errors of the calculated results compared with the real values for the center coordinates and the radius are illustrated in Fig. 2 to Fig. 4.

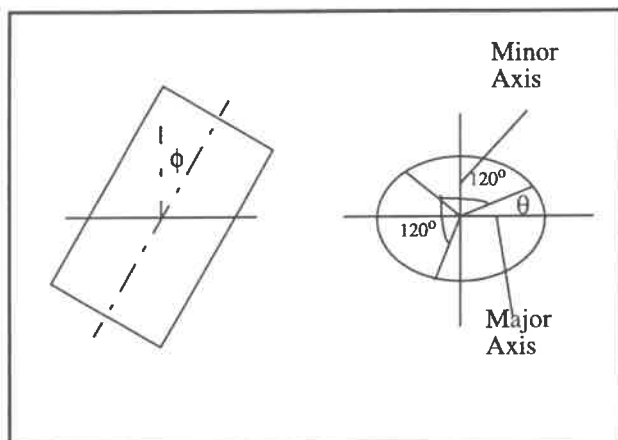


FIG.1 MEASUREMENT OF A CYLINDER

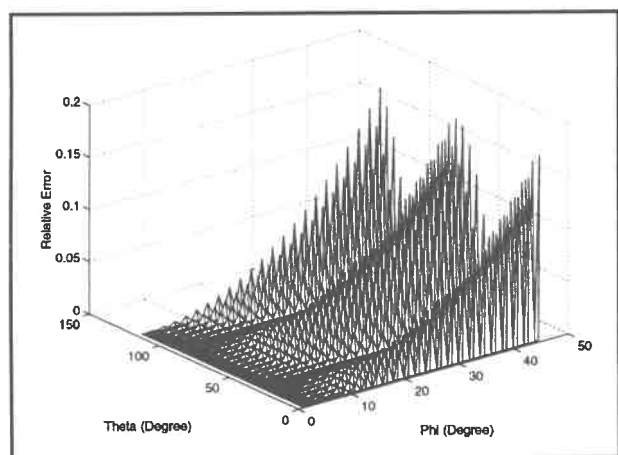


FIG.2 RELATIVE ERRORS OF THE CALCULATED X CENTER COORDINATE

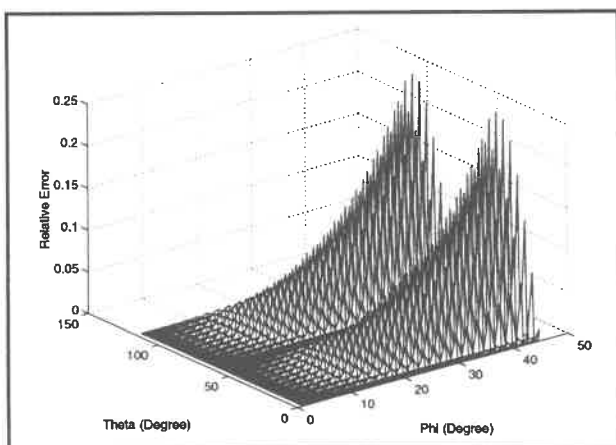


FIG.2. RELATIVE ERRORS OF THE CALCULATED Y CENTER COORDINATE

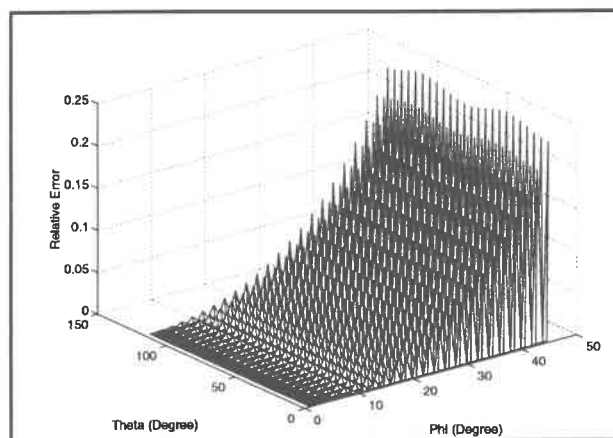


FIG. 3. RELATIVE ERRORS OF THE CALCULATED RADIUS

These significant errors are caused by the fitting algorithms only. In order to improve the calculating accuracy in this case, it is suggested that the fitting objective in each plane should be an ellipse instead of a circle. The centers of each ellipse are fitted to form the axis of the cylinder. The form error can be derived using least square method or min-max method.

ANALYSIS OF THE METHOD FOR EVALUATION OF THE ELLIPTICAL PROFILES

A commonly used method for evaluating the elliptical profile is based on the least-square method (Murthy, 1983). The general equation for an ellipse is:

$$x^2 + ay^2 + bxy + cx + dy + e = 0 \quad (1)$$

Theoretically, the center position, the major axis and the minor axis can be uniquely defined by any five points on the ellipse. In the case of multi-point measurement, the fitting objective is defined as:

$$\epsilon = \sum_{i=1}^N \left(x_i^2 + ay_i^2 + bx_iy_i + cx_i + dy_i + e \right)^2 \quad (2)$$

where: x_i, y_i : coordinates of the measurement points,

N : number of measurement points.

Differentiate eqn. (2) with respect to a, b, c, d and e :

$$\frac{\partial \epsilon}{\partial a} = 0, \frac{\partial \epsilon}{\partial b} = 0, \frac{\partial \epsilon}{\partial c} = 0, \frac{\partial \epsilon}{\partial d} = 0, \frac{\partial \epsilon}{\partial e} = 0 \quad (3)$$

From these five equations, a, b, c, d and e can be calculated and the center coordinates, major axis and minor axis can also be obtained.

Simulations were done to verify this algorithm. We define the distance between the center of the ellipse and the origin as D , the length of the half major axis as R . Ellipses with different values of D/R were simulated and fitted using the least square method. The results show that: when $D/R < 1.4142$, the relative error of the calculated results is less than 0.1%; when $D/R = 4$, the relative error is 1.797%; when

$D/R < 9.428$, the relative error is 9.337%; when $D/R = 14.142$, the relative error is 51.130%. Therefore, when D/R is greater than 9.428, the calculated results can not be accepted in engineering applications.

The reason for this phenomena is that, as D/R increases, the differences among coefficients of the eqn. (2) usually increase. Therefore, the condition number of the coefficient matrix in eqn. (3) is very high. The condition number is a measure of how close the matrix is to a singular matrix. The solutions are usually biased for a coefficient matrix which is near singular with a high condition number.

One way of improving the accuracy is to perform normalization transformation for the measured data. For the equations with the form of $AL=B$ such as eqn. (3), do the following transformation to obtain normalized equations $WQ=Y$, where:

$$\|a_i\|^2 = \sum_{j=1}^N a_{ji}^2 \quad (4)$$

$$M = \begin{bmatrix} \|a_1\| \\ \|a_1\| \\ \dots \\ \|a_1\| \end{bmatrix} \quad (5)$$

$$W = AM^{-1} \quad (6)$$

$$Y = \|B\|^{-1}B \quad (7)$$

The solutions of the normalized equation $WQ=Y$ can be used to obtain the solutions of the original equations by transformation:

$$L = \|B\|M^{-1}Q \quad (8)$$

This method is usually too involved and will reduce the speed of the software.

IMPROVED METHOD BASED ON LINEARIZATION

Linearization technique is employed to derive the improved algorithm. For the multi-point measurement process, in order to fit all the measurement points to eqn. (2), first, we select the data of five points from the measured points to plug into eqn.(2) and get the initial solution $(a_0, b_0, c_0, d_0, e_0)$. This result is usually very close to the results obtained when using all the points on the profile. Then, the right side of eqn. (2) can be expressed by the Taylor series (the terms whose power is bigger than 2 are omitted) as:

$$\begin{aligned} \varepsilon(a, b, c, d, e) &= \varepsilon(a_0, b_0, c_0, d_0, e_0) \\ &+ \left[\left(\frac{\partial \varepsilon}{\partial a} \right)_0, \left(\frac{\partial \varepsilon}{\partial b} \right)_0, \left(\frac{\partial \varepsilon}{\partial c} \right)_0, \left(\frac{\partial \varepsilon}{\partial d} \right)_0, \left(\frac{\partial \varepsilon}{\partial e} \right)_0 \right] \cdot \Delta^T \\ &+ \frac{1}{2} \Delta \cdot H \cdot \Delta^T \end{aligned} \quad (9)$$

where: $\Delta = [\Delta a, \Delta b, \Delta c, \Delta d, \Delta e]$, and $\Delta a = a - a_0$, $\Delta b = b - b_0$, $\Delta c = c - c_0$, $\Delta d = d - d_0$, $\Delta e = e - e_0$, and

$$H = \begin{bmatrix} \left(\frac{\partial^2 \varepsilon}{\partial a^2} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial a \partial b} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial a \partial c} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial a \partial d} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial a \partial e} \right)_0 \\ \left(\frac{\partial^2 \varepsilon}{\partial b \partial a} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial b^2} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial b \partial c} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial b \partial d} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial b \partial e} \right)_0 \\ \left(\frac{\partial^2 \varepsilon}{\partial c \partial a} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial c \partial b} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial c^2} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial c \partial d} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial c \partial e} \right)_0 \\ \left(\frac{\partial^2 \varepsilon}{\partial d \partial a} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial d \partial b} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial d \partial c} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial d^2} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial d \partial e} \right)_0 \\ \left(\frac{\partial^2 \varepsilon}{\partial e \partial a} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial e \partial b} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial e \partial c} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial e \partial d} \right)_0 & \left(\frac{\partial^2 \varepsilon}{\partial e^2} \right)_0 \end{bmatrix} \quad (10)$$

Differentiate eqn. (9) with respect to a, b, c, d and e , we can get:

$$H\Delta = D \quad (11)$$

where: $D = -\left[\left(\frac{\partial \varepsilon}{\partial a} \right)_0, \left(\frac{\partial \varepsilon}{\partial b} \right)_0, \left(\frac{\partial \varepsilon}{\partial c} \right)_0, \left(\frac{\partial \varepsilon}{\partial d} \right)_0, \left(\frac{\partial \varepsilon}{\partial e} \right)_0 \right]^T$

eqn. (11) can be solved to yield $\Delta a, \Delta b, \Delta c, \Delta d$ and Δe . Then $a = a_0 + \Delta a$, $b = b_0 + \Delta b$, $c = c_0 + \Delta c$, $d = d_0 + \Delta d$, $e = e_0 + \Delta e$. In order to further improve the accuracy, the obtained a, b, c, d and e can be used as the initial values and use this method to calculate again and update the most recently obtained value until the results satisfy a pre-assigned termination condition. The center coordinate of the ellipse (m, n), half major axis R and half minor axis r can be calculated by:

$$m = \frac{2ac - bd}{b^2 - 4a} \quad (12)$$

$$n = \frac{2d - bd}{b^2 - 4a} \quad (13)$$

$$R = \sqrt{\frac{2(bcd - ac^2 - d^2 + 4ae - eb^2)}{(b^2 - 4a)(1 + a - \sqrt{b^2 + (1 - a)^2})}} \quad (14)$$

$$r = \sqrt{\frac{2(bcd - ac^2 - d^2 + 4ae - eb^2)}{(b^2 - 4a)(1 + a + \sqrt{b^2 + (1 - a)^2})}} \quad (15)$$

In order to test this algorithm, ellipses with different D/R values are simulated and evaluated using the improved method. The results showed that when $D/R < 9.428$, the relative error is less than 0.930%; when $D/R = 14.142$, the relative error is 9.030%; when $D/R = 18.562$, the relative error is 23.310%. The comparison of the relative calculation error of the commonly used least square method and the improved method is shown in fig.5. It is evident that the improved method can accurately get the results in a wider range of D/R values than the commonly used least square method.

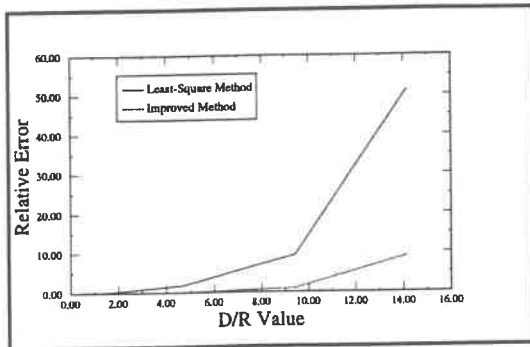


FIG. 4 COMPARISON OF THE METHOD

ROBUST ALGORITHM IN THE EVALUATION OF AN ELLIPSE

Although the improved method has good accuracy for a wide range of D/R values, it still can not guarantee that the results are always accurate. An iterative method is proposed.

Eqns. (3) can be represented as:

$$AL = B \quad (16)$$

where: $L=[a, b, c, d, e]^T$,

$$A = \begin{bmatrix} \sum_{i=1}^N y_i^4 & \sum_{i=1}^N x_i y_i^3 & \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N y_i^2 \\ \sum_{i=1}^N x_i y_i^3 & \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N x_i y_i \\ \sum_{i=1}^N x_i y_i^2 & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N x_i \\ \sum_{i=1}^N y_i^3 & \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N y_i^2 & \sum_{i=1}^N y_i \\ \sum_{i=1}^N y_i^2 & \sum_{i=1}^N x_i y_i & \sum_{i=1}^N x_i & \sum_{i=1}^N y_i & N \end{bmatrix}$$

$$B = - \begin{bmatrix} \sum_{i=1}^N x_i^2 y_i^2 & \sum_{i=1}^N x_i^3 y_i & \sum_{i=1}^N x_i^3 & \sum_{i=1}^N x_i^2 y_i & \sum_{i=1}^N x_i^2 \end{bmatrix}$$

N is the number of measurement points.

Eqn. (16) can be solved and yield a, b, c, d and e . These values can be plugged in eqns. (12) ~ (15) to get the center coordinate (m_0, n_0) , half major axis R_0 and half minor axis r_0 . Then, perform transformation to the measurement data using: $x_i = x_i - m_0, y_i = y_i - n_0, (i=1 \sim N)$. Put the transformed data into eqn. (16) to calculate the new center coordinate, half major axis and half minor axis. The process proceeds until a pre-assigned termination condition is satisfied. The simulation results showed that this method can always yield the accurate results for arbitrary D/R value. To make the process simple, the first run of calculation can be done by using the

coordinates of only five measurement points. Experiences show that generally 1~2 iterations are needed for the calculation to terminate.

The algorithm proposed can be used in the evaluation of the cylinder measured by CMMs when the axis of the cylinder is not parallel to the Z axis of the machine.

CONCLUSION

The software of the CMMs might have significant errors to make the reported results biased. The errors involved in the measurement of a cylinder by CMMs using the conventional method are demonstrated by simulation. It is argued that fitting an ellipse to the measurement data for one layer instead of a circle will yield more accurate results. The ratio D/R , which D is the distance between the center of the ellipse and the origin and R is the half major axis of the ellipse, is identified as a critical value to make the algorithm not robust. An improved algorithm based on the linearization is proposed. This algorithm shows advantages over the commonly used least square method. Furthermore, a robust algorithm is proposed which is based on iteration. These algorithms can easily be applied in the software of the CMMs.

REFERENCES

- Hockon, R. J., Raja, J. and Babu, U. (1993), Sampling Issues in Coordinate Metrology. International Forum on Dimensional Tolerancing and Metrology, CRTD-Vol. 27, 97-111. ASME.
- Hopp, T. H., (1993), Computational Metrology. International Forum on Dimensional Tolerancing and Metrology, CRTD-Vol. 27, 207-217. ASME.
- Lotze, W. (1994), General Solution for Tsebychev Approximation of Form Elements in Coordinate Measurement. Measurement, Vol. 12, 337-344.
- Murthy, T. S. R., Reddy, G. C., and Radhakrishnan, V., (1983), A Method for Evaluation of Elliptical Profiles. Precision Engineering, Vol. 5, No. 2., 77-81.
- Murthy, T. S. R., (1982), A Comparison of Different Algorithms for Cylindricity Evaluation. Int J. Mach. Tool Des. Res. Vol. 22, No.4, 283-292.
- Murthy, T. S. R., (1980), Minimum Zone Evaluation of Surfaces. Int J. Mach. Tool Des. Res. Vol. 20, 123-136. —
- Phillips, S. D., et al., (1993), Measurement Uncertainty Considerations for Coordinate Measuring Machines. NISTIR 5170.
- Shunmugam, M. S., (1987), Comparison of Linear and Normal Deviations of Forms of Engineering Surfaces. Precision Engineering, Vol. 9, No.2, 96-102.
- Shunmugam, M. S., (1991), Criteria for Computer-Aided Form Evaluation. Transactions of the ASME, Journal of Engineering for Industry, Vol. 113, 233-238.
- Waldele, et al., (1993), Testing of Coordinate Measuring Machine Software. Precision Engineering, Vol. 15, 121-123.

INVESTIGATION OF ERROR MOTIONS OF AEROSTATIC SPINDLES

M. Weck, M. F. Soltani

Fraunhofer Institut für Produktionstechnologie, IPT
Aachen, Germany

1 Introduction

The perfect function of an air-bearing is among other aspects secured through the form accuracy of the bearing. Should high guidance accuracy and stiffness be achieved, the bearing gap height and the form deviation must be kept very small. High requirements on measurement and form accuracy of the bearing parts are the main cause of high manufacturing costs of aerostatic spindle systems. To minimise manufacturing costs it is very important to know which influence the form error of rotor and stator has on the error motion of the spindle. This report describes influences on the error motion of an aerostatic spindle and gives some experimental results.

2 Theoretical model of influences on the error motion of aerostatic spindles

Error motions of aerostatic spindles are affected by the bearing configuration chosen and the manufacturing tolerances. The bearing gap is formed by the machined surfaces of the rotor and the stator of the spindle. Any inaccuracies in form and position of these surfaces lead to local pressure variations in the bearing gap. When the pressure distribution changes while the spindle is rotating, error motion occurs. In addition to the manufacturing and assembling errors, the bearing specifications, gravitational, centrifugal and other forces in a fixed direction (e.g. cutting force) as well as the spindle speed may influence the error motion of an aerostatic spindle [1]. For a better understanding of the influences mentioned above a simple model (spring model) has been developed (fig. 1). An equation for the simplified calculation of the error motion of an aerostatic journal bearing with orifice restrictors has been performed. Every pressure pad has been simplified by a spring with a constant stiffness.

The error motion in radial x and y can be written as:

$$\Delta X(\varphi_z) = \frac{1}{\sum_{i=1}^I (K_i \cdot \cos^2 \theta_i)} * \left[\sum_{i=1}^I [K_i \cdot \cos \theta_i * (\sum_{n=1}^N (B_n \cdot \cos(n \cdot \theta_i - \Delta \theta_n)) - \sum_{j=1}^J (A_j \cdot \cos(j \cdot \theta_i - j \cdot \varphi_z - \Delta \phi_j)))] + F_{z_x} + F_{ce_x} + F_{g_x} \right]$$

$$\Delta Y(\varphi_z) = \frac{1}{\sum_{i=1}^I (K_i \cdot \sin^2 \theta_i)} * \left[\sum_{i=1}^I [K_i \cdot \sin \theta_i * (\sum_{n=1}^N (B_n \cdot \cos(n \cdot \theta_i - \Delta \theta_n)) - \sum_{j=1}^J (A_j \cdot \cos(j \cdot \theta_i - j \cdot \varphi_z - \Delta \phi_j)))] + F_{z_y} + F_{ce_y} + F_{g_y} \right]$$

The variables are given in the nomenclature.

This model is an easy to handle tool which gives users the possibility to determine the radial error motion of an aerostatic spindle. Because of the approximation which are made in this model the results have to be understood also as approximations. With the help of this model statements about runout characteristics of a rotor supported by an externally pressurized gas journal bearing with special attention to the relationship between runout and manufacturing errors such as out-of-roundness of rotor and stator can be performed.

The most important statements are:

- if the rotor roundness is perfect there will be no error motion
- the stator roundness has no influence on the error motion
- the effect of a combination of a given number of supply holes (orifices) and a number of lobes on the rotor is either a "zero error" or the complete rotor error is reproduced as an error motion of the spindle [1]. For a given number of orifices l the only lobes U to contribute to an error motion are:
$$U = K \cdot l \pm 1, \text{ where } K = 1, 2, 3, \dots$$
- the frequency of the error motion is the same as the frequency of the roundness error of the rotor

3 Experimental investigation of the error motion of an aerostatic spindle

The aim of this experiment is to study the influence of inlet pressure, number of orifices and form error of the rotor on the error motion of the spindle. Two rotors with different form errors were used. The first rotor will be used as a reference. From there its form accuracy during finish machining was not deliberately affected. To study the additional influence of form errors of the rotor on the motion accuracy of the rotating axis the second rotor was manufactured with a defined roundness deviation.

Figure 2 shows a cross section of the test spindle. Up to 12 pocketed orifices can be used for the radial guidance of the rotor. Every orifice is separately supplied with pressurized air. The axial guidance of the rotor is taken over by 3 axial orifices, which are equally distributed over the circumference of the lower part of the stator. The radial error measurements were carried out with two ADE capacitance sensors, which measured against a cylindrical master. The error is calculated using the "difference method" $(A-B)/2$. This method reduces the influence of noise on the total error and eliminates the even numbered form deviations of the master. To increase the measurement accuracy evaluation of the measured values is carried out over 16 revolutions.

3.1 Experimental results by using a "perfect" rotor

The following measurements are carried out at low speed (average 50 rpm) and free run of the spindle to prevent influences from drive and unbalance on the error motion. Several measurements were performed with different supply pressure values (3, 5 and 7 bar) and number of orifices (3, 4, 6 and 12).

Figure 3 shows that the total error was more dependent on the number of orifices than on the supply pressure. There is an averaging effect only with 4, 6 and 12 orifices. The averaging effect means that the rotating accuracy is much better than the magnitude of machining errors at rotor and bearing, which were measured to about $0.23 \mu\text{m}$ and $0.31 \mu\text{m}$ respectively.

The synchronous error is a repeatable error and gives more information about the effects of systematic error sources - such as shaft form - on the error motion of the spindle. Therefore only this error type will be considered in the investigation of the influence of shaft form on rotation accuracy of the spindle.

The spectrum of the synchronous error (fig. 4) confirms the $k \cdot l \pm 1$ expectations from the spring model for $l=3, 4$ and 6 orifices. For 12 orifices the second harmonic is the largest, but the 11^{th} and 13^{th} are larger than the 12^{th} harmonic. The relation between the number of orifices and the frequency of the form deviation of the rotor will be further investigated in the next chapter.

3.2 Experimental results by using a rotor with a defined roundness deviation

The rotor used in this experiment has 11 lobes at the circumference with an amplitude of $0.5 \mu\text{m}$ (fig. 3.4). According to the spring model this shaft shape should lead to an error motion with a frequency of 11 per revolution independent from the number of orifices and the inlet pressure. The radial rotor shape was manufactured using a Fast Tool Servo System as described in [1].

The frequency analysis of the synchronous error with different number of orifices and different values for inlet pressure are shown in figure 6. At 7 and 5 bar inlet pressure the frequency $11/R$ (Revolution) could be found. With 3 orifices the $2/R$ and the $4/R$ are the dominant error frequencies. With 4 orifices the $3/R$ and $5/R$ and with 6 orifices the $5/R$ and the $7/R$ are the dominant error frequencies.

At 3 bar inlet pressure the amplitude of the 11/R increases and at 1 bar pressure the error with 11/R frequency is the dominant one and its amplitude was measured to about 0.5 μm . As the air supply pressure decreases, the correlation between theory and experimental results increases.

One reason for this behaviour can be found in the influence of gap variation and eccentricity between rotor and stator. The theory of the spring model neglects stiffness variations at each orifice. The influence of gap variation on the stiffness can be studied in figure 7, which shows the theoretical stiffness of an orifice bearing plotted against gap height at different values of inlet pressure [2]. The real gap in this case was measured to about 16 μm . At 7 and 5 bar inlet pressure a deviation of the gap height leads to high changes of stiffness from orifice to orifice. In this case the main condition of stiffness stability at each orifice in the spring model is not true. As the inlet pressure decreases, the stiffness variation because of gap height deviation decreases too. At 1 bar inlet pressure the stiffness at each orifice remains constant over a large interval of gap height. In this case the influence of the gap variation on the stiffness is neglectable and the spring model is valid.

4 Conclusion

The first step in the error motion study was to develop an approximation theory for the dependence of error motions on different influences, especially the manufacturing tolerances. Because of the very high number of parameters which influence error motions, the research work was focused on a journal bearing with pocketed orifice restrictors. With the help of the developed "spring model" it is possible to predict error motions. Several error motion measurements were performed using a rotor with "perfect form" and a second one with well known form deviations (11 UPR). The error motion was detected to be more dependent on the number of orifices than on the inlet pressure.

With the "perfect rotor" it was found out that there is a clear relationship between synchronous error frequency and number of orifices, which confirms the $K \cdot l \pm 1$ expectation from the spring model. The validity of the spring model using the rotor with 11 UPR was only proved at low values of inlet pressure (≤ 3 bar). The reason for this behaviour could be found in the fact that stiffness variations due to gap variations are disregarded in the spring model. In this case a constant stiffness over large gap variation is only true for low inlet pressure.

5 Nomenclature

l	: number of orifices	i	: orifice number
J	: maximum number of rotor lobes/ 2π	N	: maximum number of stator lobes/ 2π
K_i	: stiffness of pressure pad i	θ_i	: angle position of orifice i
ϕ_z	: angle of rotation		
F_z	: force; fixed direction	F_{ce}	: centrifugal force
F_g	: gravitational force	F_i	: inner force at orifice i
h	: theoretical gap height	h_i	: gap height at orifice i
(x,y)	: fixed coordinate system	(ξ,η)	: rotating coordinate system
B_n	: amplitude of the roundness profile of the stator with n lobes/ 2π		
A_j	: amplitude of the roundness profile of the rotor with j lobes/ 2π		
$\Delta\theta_n$	: phase of the roundness profile with n lobes/ 2π (stator)		
$\Delta\phi_j$	: phase of the roundness profile with j lobes/ 2π (rotor)		

6 References

- [1] Final Report on BRITE-EURAM Project BE-519
"High Precision High-Speed Spindle System", 1/1996
- [2] M. Weck
Werkzeugmaschinen, Fertigungssysteme, Band 2, VDI-Verlag, Düsseldorf 1991

7 Figures

The diagram illustrates a rotor-stator system. A central rotor is surrounded by a stator. Key components and labels include:

- Forces:** F_g (gravitational force), F_z (vertical force), F_i (inward radial force), F_{ce} (centrifugal force), K_i (spring force), and P_{si} (stator pressure).
- Displacements and Angles:** u (radial displacement), h (gap height), h_i (gap height at orifice i), θ_i (orifice angle), ψ_z (vertical angle), and η (rotational angle).
- Geometric Features:** Orifice i , Stator, Rotor, and Orifice i .
- Legend:** ideal (represented by a dashed line) and real (represented by a solid line).



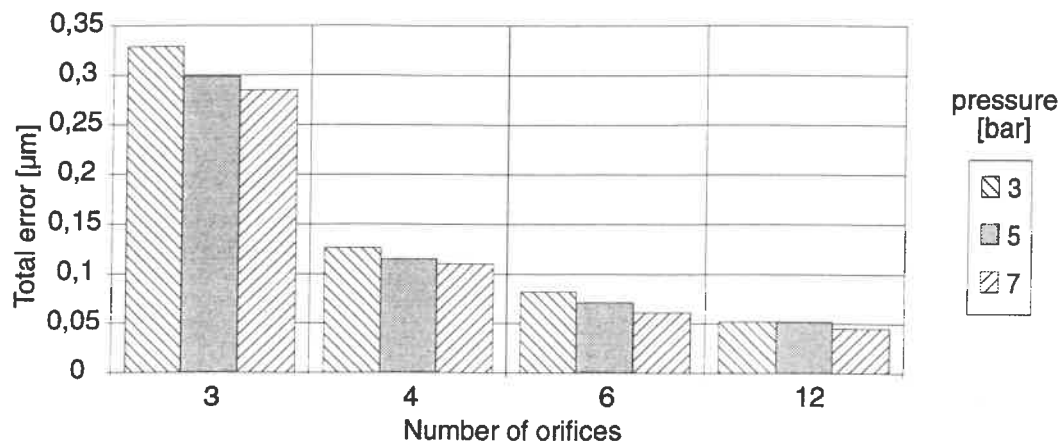


Fig. 3 Total error depending on number of orifices and inlet pressure

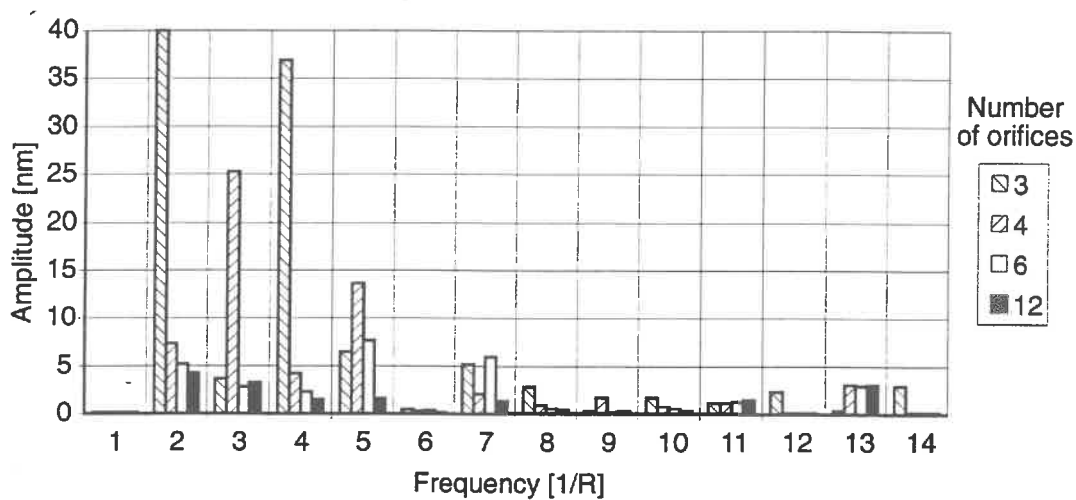


Fig. 4 Influence of the number of orifices on the amplitude of the synchronous error

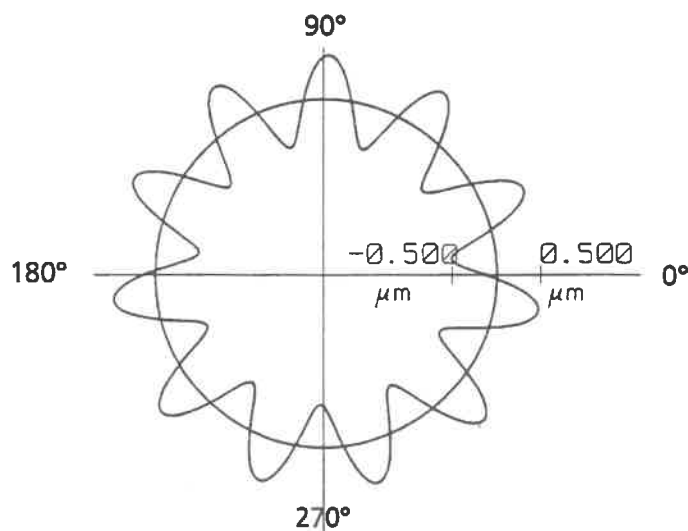


Fig. 5 Rotor with 11 undulations per revolution (UPR)

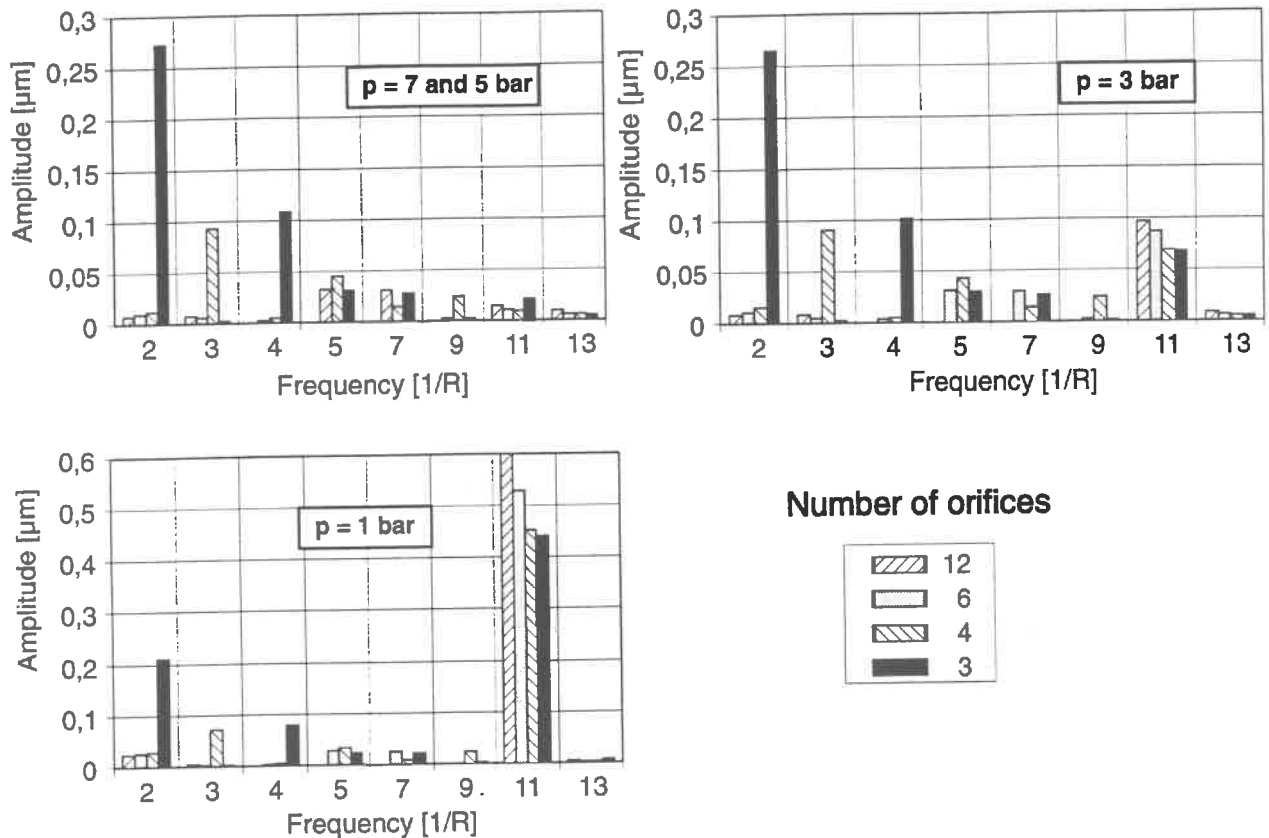


Fig. 6 Frequency analysis of the synchronous error at different number of orifices and inlet pressures (rotor with 11 UPR)

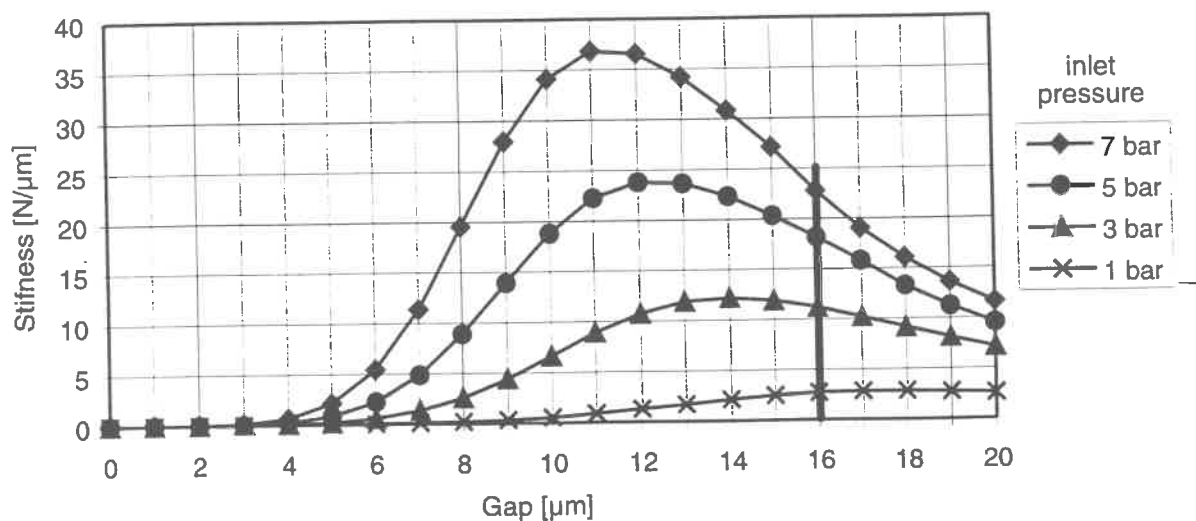


Fig. 7 Influence of inlet pressure and gap height on the Stiffness of an orifice bearing

SURFACE APPROXIMATION AND GENERATION OF COMPLEX GEOMETRIC MODEL FROM LASER SCAN DATA

Hong-Tzong Yau and Jenq-Shyong Chen
Department of Mechanical Engineering
National Chung Cheng University
Chia-Yi, Taiwan

Introduction

With the integration of computer controlled measuring machines and CAD/CAM systems, reverse engineering has emerged as an important design and manufacturing tool in recent years. Traditionally a design drawing or CAD model is created prior to the manufactured part. However, in some cases where the existing parts are modified due to poor performance or manufacturing limits, or in other cases where the design model simply does not exist, reverse engineering would then become indispensable. The goal of reverse engineering is to capture or recreate part geometry directly from measurement. Therefore, reverse engineering is an important technology aiming at closing the loop from measurement to design and manufacturing.

This paper describes the development of an intelligent reverse engineering system for automatic generation of complex geometry from laser scan data. The system consists of three major modules: (1) intelligent surface boundary detection, (2) NURBS surface approximation of scattered point data, and (3) automatic surface connection and B-Rep solid creation. The first module deals with the automatic detection and division of individual surface regions based on surface continuity considerations. Edges with curvature discontinuity will be automatically detected from the dense point data. With the grouping of various closed surface regions, planning is made to segment the entire point data map to individual surface patches. The second module then takes over by fitting a NURBS surface to each individual segmented region. First, the scattered data in the region is filtered and gridded if necessary. Then, NURBS surface approximation is made possible by minimizing the sum of squared deviations from the points to the surface. The third module matches the surface boundaries according to the position or tangency conditions based on the edge detection. The result is a complex B-Rep solid model that can be output by IGES translation to various CAD/CAM systems. The developed technology has been successfully applied to the modeling of dies and molds from automotive industry.

Feature Extraction and Edge Detection

Feature extraction here is referred to the detection and separation of geometric features from the scanned data. It consists of extraction of geometric features such as circle, plane, cylinder, etc., and detection of boundary edges that separate different surface regions. In this paper, the

emphasis has been placed on the automatic detection of surface boundaries so that various surface regions can be patched and connected for the creation of the entire complex model.

By laser scanning, objects are usually scanned by parallel scan lines. The distribution of the scanned data points typically do not follow the geometric pattern of the objects. Therefore, to segment the scanned data based on the geometric pattern, surface boundaries or edges need to be detected from the original parallel scan lines. To achieve this purpose, it is defined that points which are discontinuous in curvature be connected as "characteristic curves" and used as boundaries to separate surface regions. Any curve in a 3D space parameterized by the arc length s can be represented as

$$\vec{r}(s) = x(s)\vec{i} + y(s)\vec{j} + z(s)\vec{k} \quad (1)$$

The curvature value can be obtained by differentiating the vector against the arc length s twice: $k = |\vec{r}''(s)|$. Given three continuous points P_{i-1} , P_i , P_{i+1} , the curvature value can be approximated by

$$k = |r_i''| = \sqrt{\frac{2|r_{i+1} - r_i|}{(h_i + h_{i+1})}} \quad (2)$$

where $r_i = \left(\frac{x_i - x_{i-1}}{h_i}, \frac{y_i - y_{i-1}}{h_i}, \frac{z_i - z_{i-1}}{h_i} \right)$ and h_i is the distance between P_i and P_{i+1} .

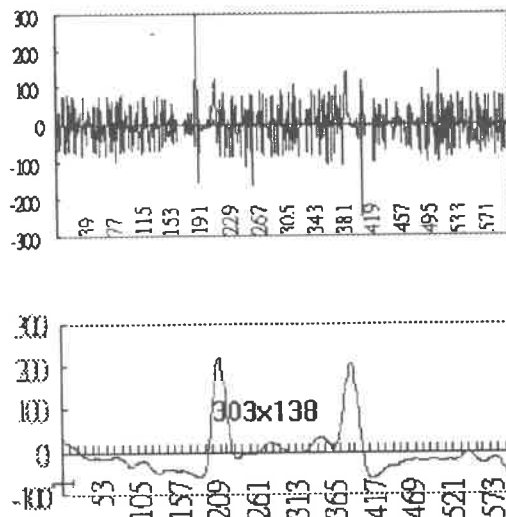


Figure1. Applying Gaussian filter to curvature data.

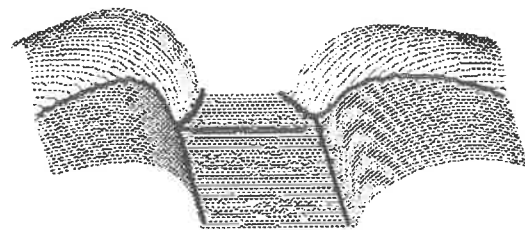


Figure2. Automatic edge detection of laser scanned data.

However, the laser scanned data is typically contaminated with substantial measurement noises. Results from direct computation of curvatures by the scanned data is often difficult to

judge and use. Therefore, Gaussian filtering is applied to the curvature data to remove high frequency noises (see Figure 1). Thus, the detection of surface boundaries can be more accurate. Figure 2 shows a successful example of finding the surface edges from a set of scanned data.

Surface Approximation

Point-to-surface conversion is the core process in reverse engineering. Since from a hierarchical planning point of view, lower level entities (points) should first be grouped into higher level entities (curves or surfaces) to make the task (creating a complex geometric model) more manageable. NURBS (Non-Uniform Rational B-Splines) has recently become the standard for surface representation in most CAD/CAM systems. NURBS is appealing in that it can represent both analytic geometry (such as conics) and free-form surfaces. In addition, NURBS has more flexibility than B-splines because it has more manipulation tools such as weights. In this work, a new approach to the fitting of NURBS is adopted [1]. Given a set of scanned data, $Q_i, i = 1, \dots, m$, and objective function to be minimized is formulated as

$$F = \sum_{j=1}^m e_j^2 \quad (3)$$

where the error

$$e_j = Q_j - \frac{\sum_{i=1}^n w_i P_i N_i(u)}{\sum_{i=1}^n w_i N_i(u)} \quad (4)$$

and w_i , P_i , and $N_i(u)$ are weights, control points, and blending functions respectively. The fitting algorithm can be summarized as [1]

1. Initialization:

Set $k = 1$

$f_k = 0$

Let $w_i = 1, i = 1, \dots, n$

$u_1 = 0$.

$$u_j = \frac{\sum_{i=2}^j |Q_i - Q_{i-1}|}{\sum_{i=2}^m |Q_i - Q_{i-1}|}, j = 2, \dots, m$$

2. Update [WP]:

$$[WP] = ([N]^T [N])^{-1} [N]^T [QNW]$$

3. Update [W]:

$$[W] = ([Q^x N]^T [Q^x N] + [Q^y N]^T [Q^y N] + [Q^z N]^T [Q^z N])^{-1} \cdot ([Q^x N]^T [NWP]^x + [Q^y N]^T [NWP]^y + [Q^z N]^T [NWP]^z)$$

4. Update [U] by solving

$$\frac{d}{du_j} |Q_j - C(u_j)|^2 = 0 \quad j = 1, \dots, m$$

5. $k = k + 1$

$$\text{Let } f_k = \left(\sum_{j=1}^m c_j^2 / m \right)^{1/2}$$

if $(f_k - f_{k+1}) < \delta$ or $k > \max_k$ go to 7

6. Go to 2

$$7. P_i = (wP)_i / w_i \quad i = 1, \dots, n$$

8. Stop

The above curve fitting algorithm can be extended to the surface case. Or the popular skinning process [2] can be used to loft the curves to the desired surface. Figure 3 shows the resulting surfaces from the data in Figure 2.

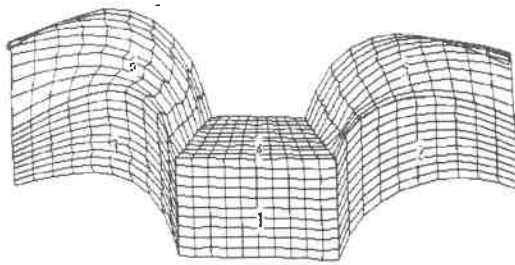


Figure3. NURBS surface approximation.

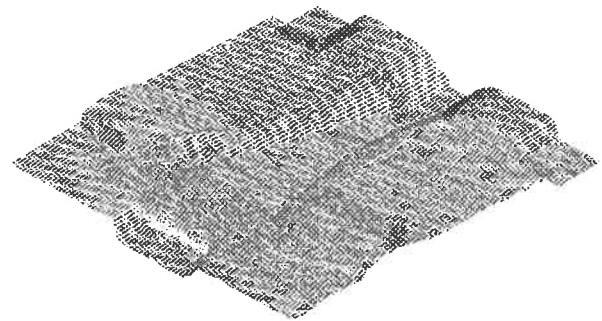


Figure4. Digitized data of a complex die Geometry.

Surface Merging and Creation of Complex Model

In the segmentation of scanned data, cases may occur that gaps or overlaps exist between surfaces. Since our objective is to create an enclosed volume for a solid model, the boundary continuity has to be maintained. Therefore, boundary merging is needed after individual surfaces

are created. In this work, merging is applied to boundaries between two surfaces to obtain C^0 and C^1 continuity.

If surfaces $S1$ and $S2$ are adjacent to each other, to merge $S2$ to $S1$ the following operations can be carried out.

$$C^0: (\bar{P}_{i,l}^w)_{S2} = (\bar{P}_{i,l}^w)_{S1} \quad (i = 1, \dots, n) \quad (5)$$

$$C^1: (\bar{P}_{i,2}^w)_{S2} = 2(\bar{P}_{i,l}^w)_{S1} - (\bar{P}_{i,l-1}^w)_{S1} \quad (i = 1, \dots, n) \quad (6)$$

Figure 4 shows a set of digitized data for a die. Using the procedures proposed in this paper, a complex geometric model consisting of 32 NURBS surfaces was created (Figure 5). Note that the surface boundaries are automatically detected by the curvature analysis, and the surface parametric distribution are generally following the geometric pattern. Figure 6 shows the shaded image.

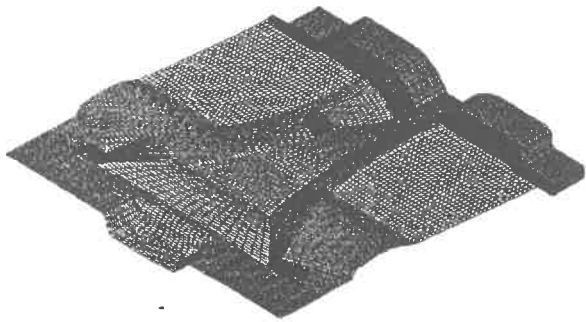


Figure5. Creation of complex surface model.

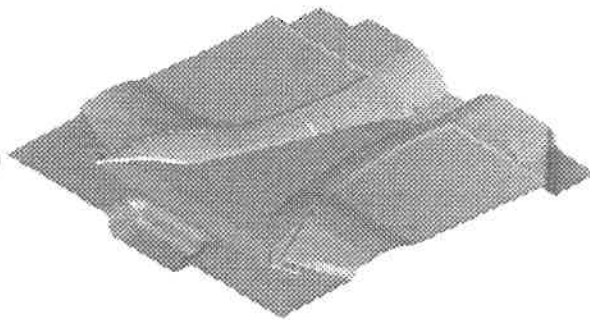


Figure6. Shaded image of the complex die surface.

References

- [1] H.T. Yau, 1996, "Reverse Engineering of Engine Intake Ports by Digitization and Surface Approximation", Int. J. of Machine Tools and Manufacture, in press.
- [2] C.D. Woodward, 1987, "Cross-Sectional Design of B-Spline Surfaces", Computers and Graphics, Vol. 11, No. 2.

MACHINE TOOL ACCURACY CHARACTERIZATION USING TELESCOPIC BALLBAR

A. Chukwujekwu Okafor and Yalcin M. Ertekin
Laboratory for Industrial Automation and Flexible Manufacturing
Department of Mechanical and Aerospace Engineering and Engineering Mechanics
University of Missouri-Rolla
Rolla, MO 65409-0050

ABSTRACT

The increasing global market competition has focused the attention of manufacturers on how to manufacture products at the least cost and in the shortest time with the highest precision to increase manufacturing efficiency, productivity, and improved quality [1]. Increasing demand for both higher machine tool accuracy and machining tolerances has resulted in a great need to measure and characterize the accuracy of NC machine tools and minimize machine tool induced errors on machined components before cutting parts, instead of inferring it from parts already machined. Statistical process control (SPC) provides no more than after-the-fact deduction of process capability.

This paper presents our on-going research on machine tool accuracy characterization and error compensation using telescopic magnetic ballbar. The results of the accuracy/repeatability assessment of an old Bridgeport Series I CNC 2 1/2 axes vertical machining center are presented. The effects of speeds and feeds are also presented.

INTRODUCTION

Most of the accuracy characteristics of machine tools deteriorate as a function of running time due to wear of tool and machine components [2]. Accuracy also decreases after any job involving unusually high cutting forces, whenever the machine tool is moved, after any spindle crash into the workpiece and also due to environmental influence. Any machine tool which is out of alignment operates under less than optimum conditions, resulting in higher maintenance, part and labor costs, scrap and re-work rates. Out of alignment cutting tools will wear out faster under increased stresses and break down more often, thus both productivity and operational profits will be reduced [3].

There are three main sources of errors in machine tools that determine machine tool accuracy. 1) Errors due to geometric inaccuracies which are caused by the mechanical imperfections of the machine tool structure and the misalignment of machine tool elements. They all change gradually due to wear of the components. The effect of the geometrical inaccuracies is to produce errors in the squareness and parallelism between the machine's moving elements. 2) Thermally induced errors are generated by environmental temperature and local sources of heat from drive motors, gear trains and other transmission devices, thus causing expansion, contraction and deformation of the machine tool structure and resulting in positional errors between the cutting tool and workpiece. The machine tool elements particularly effected by self generated thermal distortion are spindles and leadscrews. 3) Load induced error which are produced by varying mechanical loads applied to the machine tools, therefore, causing elastic strain of the machine tool structure [4].

There are various methods for assessing and correcting undesired errors in the accuracy of machine tools. These procedures are categorized mainly in two sections where the machine tools are measured and to those where machined part is measured. The methods are given as: 1) Conventional geometric measurement of machine tools using mechanical devices such as step gages, squares, levels. 2) Monitoring of test pieces. 3) SPC monitoring of the machined workpieces. 4) Measuring the machine by modern test artifacts such as telescopic Double BallBar (DBB) or laser/optical calibration. Although laser calibration procedures may be time consuming, they are complete and absolute in their error mapping test procedures. This allows the overall volumetric envelope of the machine tool to be re-calibrated by adjusting/correcting software of CNC through error table in the memory. Complementary to such absolute techniques are artifact based assessment procedures ranging from telescopic/fixed length DBB's [5, 6] to circularity tests using a circular masterpiece [7].

EXPERIMENTAL PROCEDURES

A Renishaw serial BallBar was used for data collection with a nominal length of 150 mm (including extension bar). The BallBar system had a resolution of 0.1 μm and accuracy of $\pm 1.0 \mu\text{m}$ (at 20 °C) and calibrated prior to data acquisition with a Zerodur calibrator.

A schematic diagram of the experimental set-up for data collection using BallBar is given in Figure 1. For dynamic data capturing, the Bridgeport CNC milling machine was programmed to move in the xy plane using a 100 mm circular contouring radius in both Clockwise (CW) and Counter Clockwise (CCW) directions. Angular overshoot was 180 ° before and after data capture and 360 ° data capture arc was utilized. Three different feed rates of 127, 254, and 508 mm/min and three cutting speeds of 60, 100, and 130 RPM were used to study the effects of feed rate and cutting speed. Sampling rate of the dynamic data capture software was 7.81 samples/s except for the feedrate of 127 mm/min where 1.95 samples/s was used. For static data capturing the same set-up was used. The CNC milling machine was programmed to move into 8 target points in circular interpolation both CW (G2) and CCW (G3) directions and dwelled for static data capturing in the xy plane using 100 mm radius. Feedrates were the same as for dynamic data capture but only 60 RPM was utilized. Target points are given in Figure 2.

Renishaw calibration software system was used to generate circular plots and calculate error diagnostics as well as statistical repeatability.

RESULTS AND DISCUSSION

Samples of captured dynamic circularity error plots using ISO 230-1 are shown in Figure 3 for two different feedrates of 254 and 508 mm/min and two spindle speeds of 60 and 130 RPM. ISO 230-1 circularity error value is the difference between maximum outward deviation and maximum inward deviation from the best circle through the captured data. Using different cutting speeds of 60, 100, and 130 RPM and feedrates of 254 mm/min and 508 mm/min, the resulting circularity error plots shown in Figure 4a indicates that the error is not very sensitive to spindle speed changes in both CW and CCW directions. Also this graph indicates that circularity error increases when feedrate is increased from 254 to 508 mm/min in CW direction whereas in CCW direction, there is no significant change in circularity error. Figure 4b shows similar result when circularity error is plotted as a function of feed at fixed cutting speed (60 RPM).

In another attempt to investigate feedrate effects on the circularity error of the CNC milling machine, the machine was programmed to run four runs (2 CW, 2 CCW @ 60 RPM) at feedrates of 254 and 508 mm/min and bi-directional circularity error data were captured. The circular diagnostic plots are given in Figure 5. For a feedrate of 254 mm/min, negative backlash error in the Y axis direction is the highest error with a magnitude of -19.1 μm (+Y axis), -18.2 μm (-Y axis) which constitutes 13% of the independent circularity errors. Reversal spikes and backlash errors in the X axis are second and third highest ranked errors. A Renishaw Ballbar diagnostic printouts for feedrates of 254 and 508 mm/min are given in Figure 6a and 6b respectively.

Servo mismatch is the biggest error component at the higher feedrate of 508 mm/min and followed by negative backlash errors in Y and X axes respectively (Figure 6b). High negative backlash errors are probably caused by play in the guideways of the machine causing a dwell in motion when the direction in which the machine is being driven changes. Another possible cause may be that the CNC milling machine is affected by encoder hysteresis. Servo mismatch value is the time in milliseconds by which one of the machines axis servos leads other. From the plots in Figure 5 servo mismatch is clearly seen where the plots have an oval shape along the 45° (CW) and 135° (CCW). The magnitudes of servo mismatch errors are -3.75 ms and -3.60 ms for feedrates of 254 and 508 mm/min respectively. Negative sign indicates that X axis leads Y axis in XY plane. Servo mismatch error can be corrected by balancing the loop gains of the axes servos by adjusting the machine controller. In this case X axis gain should be lowered.

Bi-directional statistical repeatability of the Bridgeport CNC milling machine is calculated according to ISO 230-2 by the diagnostic software. Using 8 target positions, static data were captured at different feedrates (127, 254, and 508 mm/min @60 RPM) with runs of 5 (3 CCW, 2CW), 8 (4CCW, 4CW), and 8 (4CCW, 4CW) respectively. Statistical and maximum repeatability values are plotted in Figure 7. The figure shows that repeatability values are higher at a feedrate of 254 mm/min and lower at 127 and 508 mm/min. All values are lower than factory specification of ± 0.0076 mm. Circular plots of captured static data are given in Figure 8, which also indicates larger errors for the feedrate of 254 mm/min.

Captured static data was also used to calculate squareness and scale mismatch errors which are plotted in Figure 9a and 9b respectively. Squareness error is the angle between x and y axes in the xy plane. A negative squareness error indicates that the angle between the two axes is less than 90°. The plot shows that squareness error is highest at a feedrate of 254 mm/min. The squareness error may be caused by worn machine guideways causing a certain amount of play in the axes or may be the axes are bent locally or there may be an overall axis misalignment in the machine. Figure of merit is the scale mismatch error normalized to units of $\mu\text{m}/\text{m}$ to give a scale mismatch error which is independent of the ballbar's length. It indicates that one of the machine's axes is either over traveling or under traveling. The graph given in Figure 9b shows that the higher the feedrate the better the scale mismatch error (figure of merit).

CONCLUSIONS AND FUTURE WORK

The results obtained from the experiments conducted with ballbar show that circular contouring errors are highly dependent on the feedrate used and not so sensitive to spindle speeds used in this research. Generally a medium feedrate of 254 mm/min consistently gave higher errors. This shows the importance of machine tool characterization in a real manufacturing environment, where one can identify less favorable cutting conditions of the machine tool and avoid using them during production. The Bridgeport CNC milling machine was diagnosed with large backlash and servomismatch errors indicating that the machine is over due for maintenance. In the future same tests will be conducted after proper maintenance.

REFERENCES:

1. Sydenham, P. H., Thorn, R., "Handbook of Measurement Science", John Wiley & Sons, 1992.
2. Torvinen, S., Andersson, P. H., Vihinen, J., Holsa, J., "An Off-line Condition Monitoring System for Machine Tools", Proceedings of LANDAMAP 95, Southampton, UK, July 1995, pp. 237-250.
3. Hamar, M., Moreyn, C., "Flat, Straight, Square, Parallel", Cutting Tool Engineering, Oct. 1993, pp. 62-69.
4. Ford, D. G., Proceedings of LANDAMAP 95, Southampton, UK, July 1995, pp. 277-286.
5. Kakino, Y., Ihara Y., Nakatsu, Y. J., Okamura, K., "Measurement of Motion Errors of NC Machine Tools & Diagnosis of Their Origins by Using Telescopic Magnetic BallBar Method", Annals of CIRP, Vol. 36, 1, 1987.
6. Knapp, W., "Test of the 3-D Uncertainty of Machine Tools & Measuring Machines and its Relation to Machined Errors", Annals of CIRP, Vol. 32, 1983, pp. 459-464.
7. Bryan, J. B., "A simple Method for Testing Measuring Machines & Machine Tools", Precision Engineering, 1982, Part1 (Vol. 4, pp. 61-69), part 2 (Vol 4, pp125-138)
8. Renishaw calibration software manual.

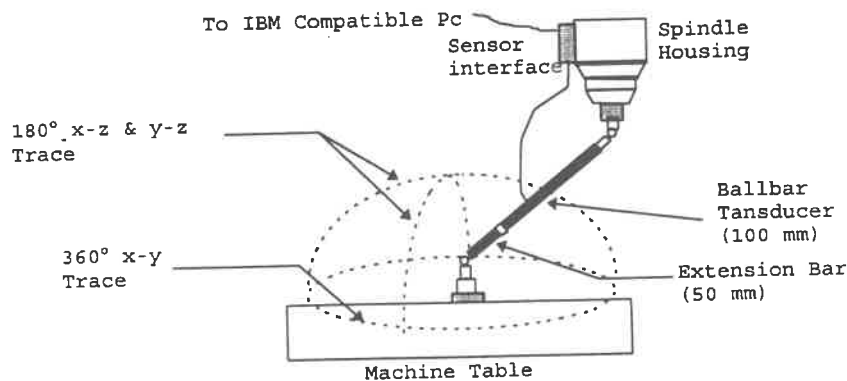


Figure 1. Experimental set-up.

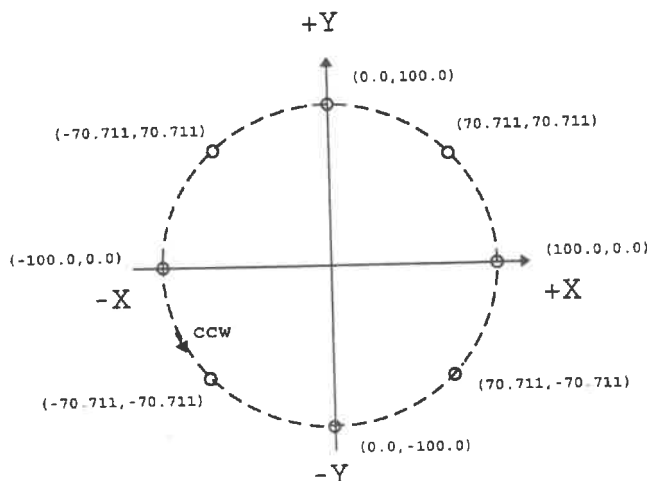
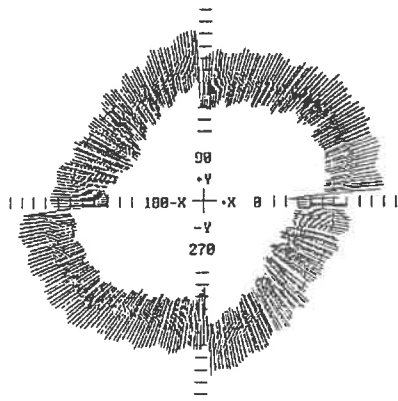
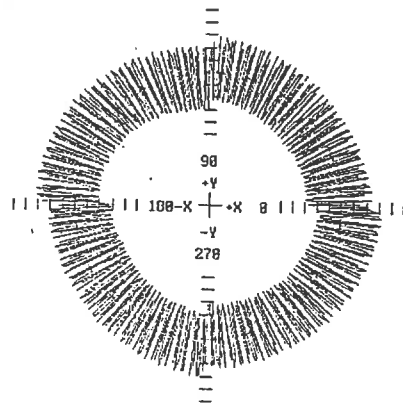


Figure 2. Static data capture target locations.



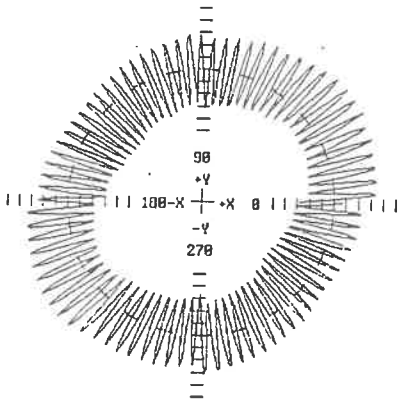
Scale: 7.8 $\mu\text{m}/\text{div}$

a) $F=254 \text{ mm/min}$, 60RPM, CW



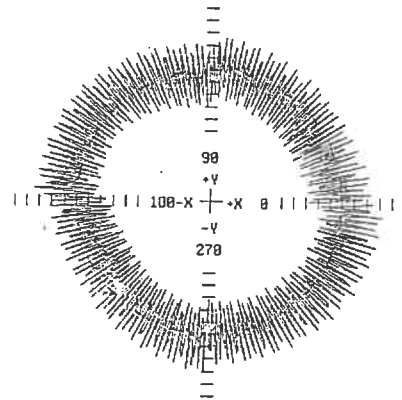
Scale: 28.8 $\mu\text{m}/\text{div}$

b) $F=254 \text{ mm/min}$, 130RPM, CCW



Scale: 28.8 $\mu\text{m}/\text{div}$

c) $F=508 \text{ mm/min}$, 60 RPM, CW



Scale: 28.8 $\mu\text{m}/\text{div}$

d) $F=508 \text{ mm/min}$, 130 RPM, CCW

Figure 3. Dynamic Circularity error plots.

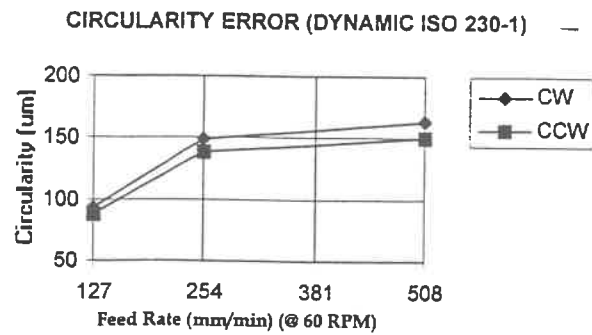
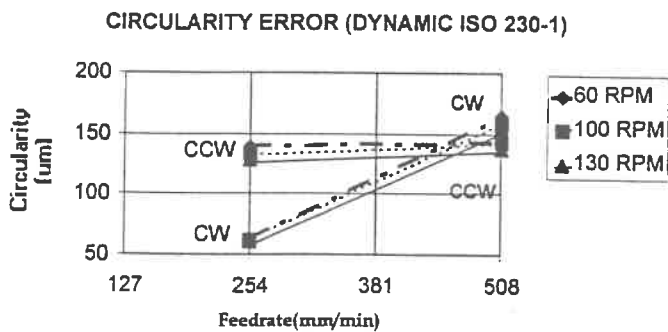
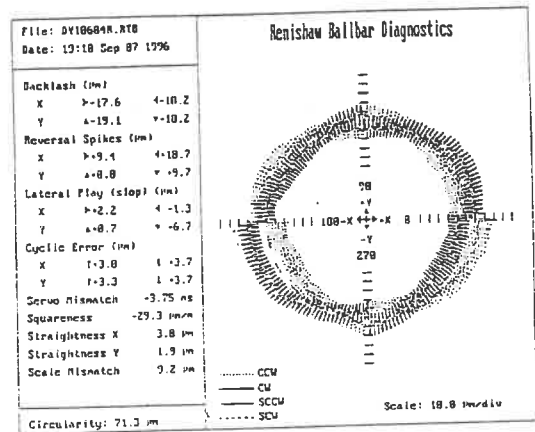
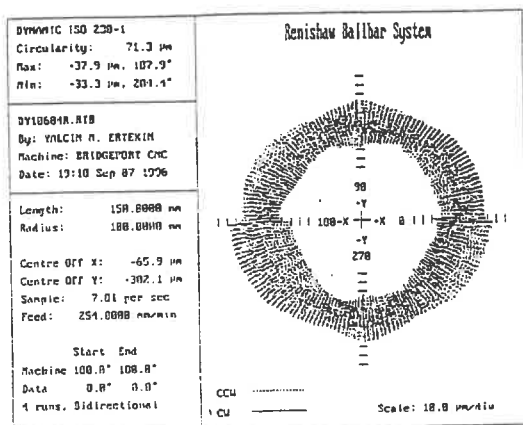
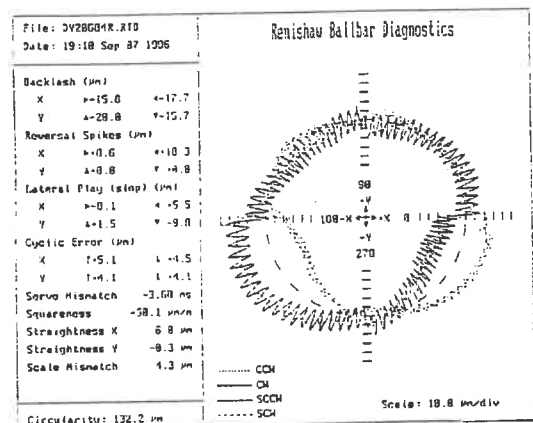
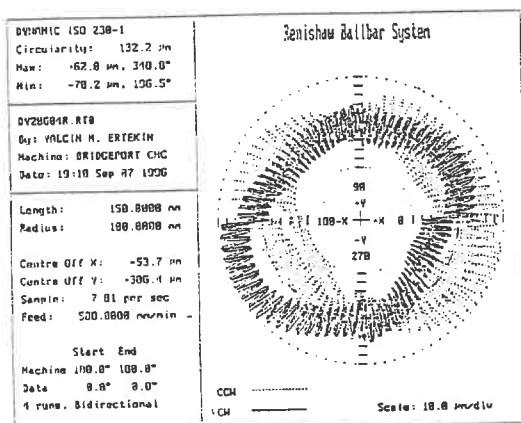


Figure 4. Circularity error.



a) F=254 mm/min, 60 RPM, 4 runs (2CW, 2CCW)



b) F=508 mm/min, 60 RPM, 4 runs (2CW, 2CCW)

Figure 5. Circular diagnostic plots.

Renishaw Ballbar Diagnostics			
ERROR	MAGNITUDE	INDEPENDENT CIRCULARITY	ADJUDICATING
Backlash X	>-17.6 <-10.2 μ m	10.2 μ m (12x)	(3)
Backlash Y	>-19.1 <-10.2 μ m	19.1 μ m (13x)	(1)
Reversal Spikes X	>-9.4 <-18.7 μ m	10.7 μ m (12x)	(2)
Reversal Spikes Y	>-8.8 <-9.7 μ m	9.7 μ m (7x)	(5)
Lateral Play (slope) X	>-2.2 <-1.3 μ m	1.2 μ m (1x)	(13)
Lateral Play (slope) Y	>-0.7 <-6.7 μ m	4.2 μ m (3x)	(7)
Cyclic Error X	>-3.0 <-3.7 μ m	1.0 μ m (1x)	(11)
Cyclic Error Y	>-3.3 <-3.7 μ m	1.9 μ m (1x)	(10)
Servo Mismatch	-3.75 ms	15.9 μ m (11x)	(4)
Squareness	-29.3 μ m/m	2.9 μ m (2x)	(8)
Straightness X	3.0 μ m	2.9 μ m (2x)	(9)
Straightness Y	1.0 μ m	1.5 μ m (1x)	(12)
Scale Mismatch	9.2 μ m	4.6 μ m (3x)	(6)
Cyclic Pitch X	2.5 μ m		
Cyclic Pitch Y	2.5 μ m		
Centre Offset X	-65.9 μ m		
Centre Offset Y	-301.4 μ m		
Circularity	71.3 μ m		

File: DY10684R.RTB
By: YALCIN M. ERTEKIN
Machine: BRIDGEPORT CNC
Date: 19:10 Sep 07 1996
Feed: 254 mm/min
Sample: 7.0125 /sec

Renishaw Ballbar Diagnostics			
ERROR	MAGNITUDE	INDEPENDENT CIRCULARITY	ADJUDICATING
Backlash X	>-15.0 <-17.7 μ m	17.7 μ m (12x)	(3)
Backlash Y	>-20.8 <-15.7 μ m	20.8 μ m (14x)	(2)
Reversal Spikes X	>-0.6 <-10.3 μ m	10.3 μ m (7x)	(4)
Reversal Spikes Y	>-0.8 <-9.8 μ m	8.8 μ m (8x)	(13)
Lateral Play (slope) X	>-0.1 <-5.5 μ m	4.3 μ m (3x)	(8)
Lateral Play (slope) Y	>-1.5 <-9.0 μ m	5.1 μ m (4x)	(5)
Cyclic Error X	>-5.1 <-4.5 μ m	2.5 μ m (2x)	(9)
Cyclic Error Y	>-4.1 <-4.1 μ m	2.8 μ m (1x)	(11)
Servo Mismatch	-3.68 ms	30.5 μ m (21x)	(1)
Squareness	-58.1 μ m/m	5.8 μ m (3x)	(6)
Straightness X	6.8 μ m	4.6 μ m (3x)	(7)
Straightness Y	-0.3 μ m	8.2 μ m (8x)	(12)
Scale Mismatch	4.3 μ m	2.2 μ m (1x)	(10)
Cyclic Pitch X	5.0 μ m		
Cyclic Pitch Y	5.0 μ m		
Centre Offset X	-55.2 μ m		
Centre Offset Y	-304.5 μ m		
Circularity	132.2 μ m		

File: DY20684R.RTB
By: YALCIN M. ERTEKIN
Machine: BRIDGEPORT CNC
Date: 19:10 Sep 07 1996
Feed: 500 mm/min
Sample: 7.8125 /sec

a) F=254 mm/min, 60 RPM, 4 runs (2CW, 2CCW)

b) F=508 mm/min, 60 RPM, 4 runs (2CW, 2CCW)

Figure 6. Renishaw ballbar diagnostic printout.

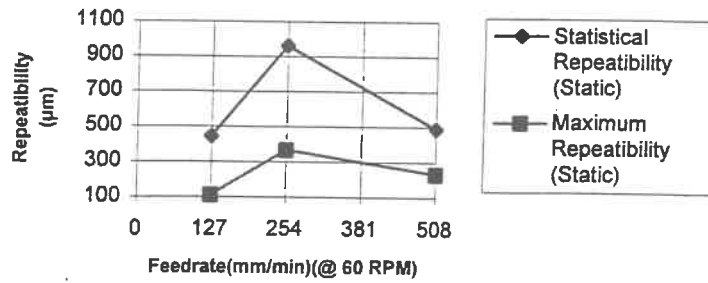
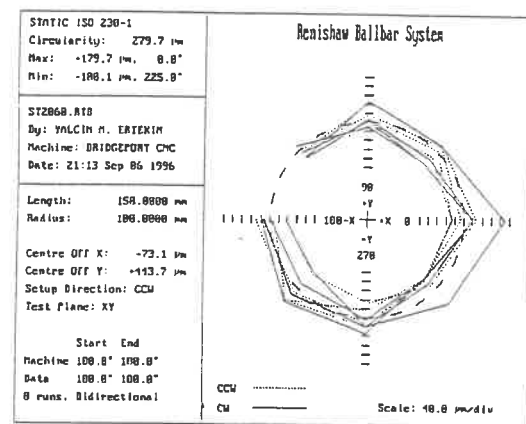
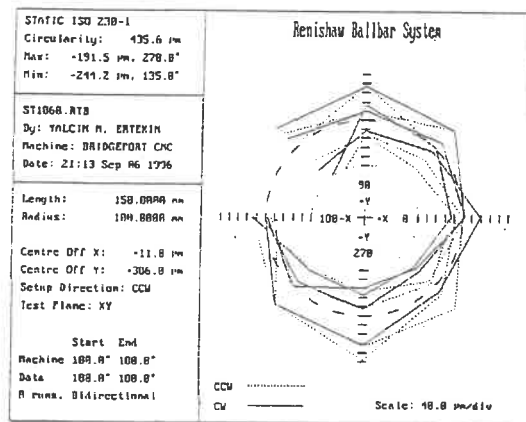


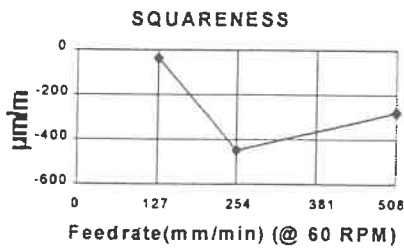
Figure 7. Statistical and maximum repeatability.



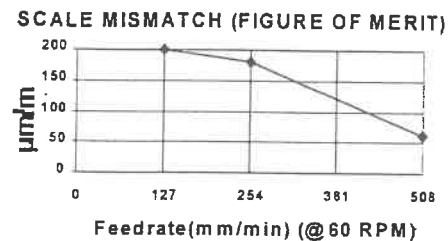
a) F=254 mm/min, 60 RPM, 8 runs (4 CW, 4 CCW)

b) F=508 mm/min, 60 RPM, 8 runs (4 CW, 4 CCW)

Figure 8. Static captured circular plots.



(a)



(b)

Figure 9. Squareness and scale mismatch errors.

A NEW TECHNIQUE FOR VOLUMETRIC ERROR ASSESSMENT OF CNC MACHINE TOOLS INCORPORATING THE BALL-BAR MEASUREMENT AND THE 3D VOLUMETRIC ERROR MODEL

Heui Jae Pahk , Young Sam Kim, Joon Hee Moon
Dept. of Mechanical Design and Production Eng.
Seoul National University
Seoul, Korea

ABSTRACT

This paper presents an useful technique for assessing the volumetric errors in multiaxis machine tools using the kinematic double ball bar. This system has been developed based on the volumetric error model which describes the 3 dimensional errors of machine tools. The developed system inputs the measured radial data performed on the 3 orthogonal planes, analyzing the parametric errors such as positional, straightness, angular, squareness, and backlash errors. The developed system also can assess dynamic performance of machine tools such as errors due to the servo gain mismatch. The developed system has been tested in a practical machine tool, and showed high potential for error assessment of multi-axis machine tools.

1. INTRODUCTION

Development of efficient techniques for performance verification of machine tools has been considered as a important task for accuracy enhancement and quality assurance for users and manufacturers of machine tools and coordinate measuring machines.

In order to assess the performance checking of machine tools, the double ball bar or kinematic ball bar has been frequently used. There are several previous research[1,2,5,3,4] on the kinematic ball bar measurement for checking accuracy of machine tools. This paper presents a new technique for analyzing the 3 dimensional parametric errors of CNC machine tools in the volumetric sense using the ball bar measurement.

2. ERROR MEASUREMENT USING THE KINEMATIC BALL BAR

Fig.1 shows a typical setup of a commercially available ball bar, where one ball is fixed on the table of machine tool and the other ball is attached to the spindle of the machine tool.

In the Fig.1, let $O(0,0,0)$ be the centre point of the ball on the table, and $P(X,Y,Z)$ be nominal coordinates of the ball centre attached to the spindle. When the machine is commanded to move to $P(X,Y,Z)$ position, the actual position of the machine is assumed $P'(X',Y',Z')$. Thus the machine geometric error can be defined as the difference between the two coordinates. That is,

$$\Delta X = X' - X, \Delta Y = Y' - Y, \Delta Z = Z' - Z \quad (1)$$

where $\Delta X, \Delta Y, \Delta Z$ are the X, Y, Z error components at the nominal position $P(X,Y,Z)$ position with respect to the $O(0,0,0)$ point. When the error components $\Delta X, \Delta Y, \Delta Z$ are present, the error in the distance between the two points, ΔR , can be evaluated as,

$$(R + \Delta R)^2 = X'^2 + Y'^2 + Z'^2 = (X + \Delta X)^2 + (Y + \Delta Y)^2 + (Z + \Delta Z)^2 \quad (2)$$

where R is the nominal distance between the two points.

When ignoring the second order terms of error components, and remembering $R^2 = X^2 + Y^2 + Z^2$

$$\Delta R = (X\Delta X + Y\Delta Y + Z\Delta Z) / R \quad (3)$$

Eq(3) gives the error in the length direction of the ball bar when the error components ($\Delta X, \Delta Y, \Delta Z$) are present during the machine movement from O to P , and it can be used for the error diagnosis of the machine tools.

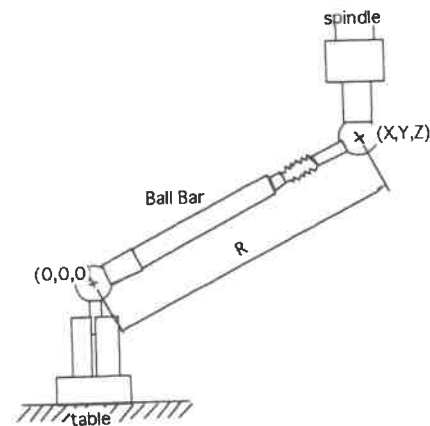


Fig.1 ERROR MEASUREMENT USING KINEMATIC BALL BAR

3. PARAMETRIC ERROR MODELING FOR MACHINE TOOLS

It is widely acknowledged that there are 6 geometric error components when a machine element moves along the guideway, and they are 3 translational

errors (positional error, horizontal/vertical straightness error) and 3 rotational errors (pitch, yaw, and roll errors). When multi-axis movement is considered, each axis has the above 6 geometric error components, and 3 squareness errors between X-Y axis, Y-Z axis, and Z-X axis. There are also dynamic errors due to servo gain mismatch between X-Y axis, Y-Z axis, and Z-X axis. In this paper, a polynomial based error modeling has been performed for the above geometric/parametric error components of machine tools, prior to the error analysis.

3.1 POSITIONAL ERROR COMPONENTS

Positional error components, $\delta x(X)$, $\delta y(Y)$, $\delta z(Z)$ along X, Y, Z axis, which are mainly due to scale errors, lead screw pitch errors, are usually defined as the difference between the actual coordinates and the nominal coordinates along the X, Y, Z axis. Thus it is possible to model the positional error as a polynomial function of position along each axis. Therefore the X positional error components are,

$$\begin{aligned}\delta x(X) &= dx_{x1} * (X/R) + dx_{x2} * (X/R)^2 + \dots \\ &= \sum_{i=1}^N dx_{xi} * (X/R)^i\end{aligned}\quad (4)$$

where X/R is the dimensionless coordinates of position along X axis, and dx_{xi} is the coefficients of the polynomial. The Y, Z positional errors can be similarly modeled.

3.2 STRAIGHTNESS ERROR COMPONENTS

Straightness error is mainly due to non-straightness of guideway and due to the bearing interfaces in machine tools. As the straightness error is defined as the perpendicular deviation along each axis, it can be modeled as the polynomial functions beginning with second order terms, while the first order polynomial function is considered for modeling the squareness error between the neighboring axis. That is,

$$\begin{aligned}\delta y(X) &= dy_{x2} * (X/R)^2 + dy_{x3} * (X/R)^3 + \dots \\ &= \sum_{i=2}^N dy_{xi} * (X/R)^i\end{aligned}\quad (5)$$

$$\begin{aligned}\delta z(X) &= dz_{x2} * (X/R)^2 + dz_{x3} * (X/R)^3 + \dots \\ &= \sum_{i=2}^N dz_{xi} * (X/R)^i\end{aligned}\quad (6)$$

where dy_{xi} , dz_{xi} are the coefficients of the polynomial function to be determined. The rest of straightness error components along Y, Z axis can be similarly modeled.

3.3 ANGULAR ERROR COMPONENTS

PITCH AND YAW ERROR

Pitch and yaw errors are the angular error components in the perpendicular direction along each axis, and they are influenced by the guideway geometry and bearing

interfaces of machine tools, thus can be related to the straightness profile along each axis. For example, pitch error, $Ex(Y)$, along Y axis can be derived as the derivative of the vertical straightness error along the y axis. That is,

$$Ex(Y) = \partial(\delta z(Y)) / \partial y = \sum_{i=2}^N i * dz_{yi} * (Y/R)^{i-1} / R \quad (7)$$

Yaw error component, $Ez(Y)$, along Y axis can be derived as the derivative of the horizontal straightness error along the Y axis, that is,

$$Ez(Y) = -\partial(\delta x(Y)) / \partial y = -\sum_{i=2}^N i * dx_{yi} * (Y/R)^{i-1} / R \quad (8)$$

where the minus sign is considered for the sign convention of the angular error components. Similarly, the rest of pitch and yaw errors are modeled as the derivatives of the respective straightness error profiles.

3.4 BACKLASH ERROR COMPONENTS

Backlash error, or reversal error, is mainly caused by the backlash in the screw/gear assembly during the motion reversal.

When the amount of backlash in the X axis is Bx , the error in the X axis due to the backlash, ΔX , can be modeled as follows.

$$\Delta X = -Bx/2 * \text{sign}(dX/dt) \quad (9)$$

where $\text{sign}()$ is sign function returning the sign of the term inside the bracket, and dX/dt is the time derivative of position, which is velocity. The minus sign in eq(9) is to give the positive backlash when the motion is changed from the forward direction to the reverse direction.

Similarly, the backlash errors in the Y, Z axis can be modeled.

3.5 SQUARENESS ERROR

Squareness error is defined as the out of squareness between the two nominally orthogonal axis, and it is mainly due to the misalignment, or misassembly in the orthogonal axes.

Let a be the amount of nonsquareness of X axis from the nominal X axis (in XY plane) at the the position of distance R from the origin point. Then the Y positional error at X position, ΔY , due to the nonsquareness of X axis is as follows,

$$\Delta Y = -a * X/R \quad (10)$$

Similarly, the rest of squareness error components can be modeled.

3.6 ERRORS DUE TO SERVO GAIN MISMATCH

When the amplifier gains of the servo drives for the axis motion are not properly matched, there exists a kind of steady state following errors. The steady state

following error, ΔR , between the actual position and the command position in the radial direction can be derived from the control theory (e.g.[7])

$$\begin{aligned}\Delta R &= M_{xy} (X/R) (Y/R) \quad \text{for CCW rotation} \\ &= -M_{xy} (X/R) (Y/R) \quad \text{for CW rotation}\end{aligned} \quad (11)$$

where M_{xy} is the coefficient for the gain mismatch between X-Y axis. Eq(8) represents the circular error which is influenced by the gain mismatch. Similarly, the circular errors due to the gain mismatch between the Y-Z axis and Z-X axis can be modeled.

4. VOLUMETRIC ERROR MODEL AND THE BALL BAR MEASUREMENT OF MACHINE TOOLS

4.1 VOLUMETRIC ERROR EQUATIONS

Three dimensional error behavior, that is, volumetric error consideration is vital for error assessment of machine tools in 3D working space. In this paper, the influence of volumetric error model on the ball bar measurement has been assessed, and appropriate error diagnosis has also been performed for a practical machine tool.

Volumetric error equations generally can be derived by the kinematic chain of each machine element when the parametric error components are calibrated (e.g.[6]), and they are dependent on the kinematic configuration of machine tools.

For the column type milling machine, the volumetric error equations[6] are,

$$\begin{aligned}\Delta X &= \delta x(X) - \delta x(Y) + \delta x(Z) \\ &\quad + Z[-E_y(Y) - b_1 - E_y(X)] + Y[E_z(Y) + E_z(X)] \\ &\quad + Yp[E_z(X) + E_z(Y) - E_z(Z)] \\ &\quad + Zp[-E_y(X) - E_y(Y) + E_y(Z)] \\ \Delta Y &= -\delta y(X) + \delta y(Y) + \delta y(Z) \\ &\quad + X[-E_z(X) + a] + Z[Ex(X) + Ex(Y) - b_2] \\ &\quad + Xp[-E_z(X) - E_z(Y) + E_z(Z)] \\ &\quad + Zp[Ex(X) + Ex(Y) - Ex(Z)] \\ \Delta Z &= -\delta z(X) - \delta z(Y) + \delta z(Z) + XE_y(X) \\ &\quad + Y[-Ex(X) - Ex(Y)] \\ &\quad + Xp[Ex(X) + Ex(Y) - E_y(Z)] \\ &\quad + Yp[-Ex(X) - Ex(Y) + Ex(Z)]\end{aligned} \quad (12)$$

4.2 INTEGRATION OF THE VOLUMETRIC ERROR EQUATIONS AND THE BALL BAR MEASUREMENT

The influence of the volumetric error behaviour of machine tools on the circular error pattern of the ball bar measurement can be considered when eq(12) are applied to the circular error equation (eq.1). Thus when the volumetric error behaviour is involved for the the column type milling machine, the circular error of ball bar measurement is,

$$\Delta R = (X \Delta X + Y \Delta Y + Z \Delta Z) / R$$

$$\begin{aligned}&= \Delta X(X/R) + \Delta Y(Y/R) + \Delta Z(Z/R) \\ &= dxx_1(X/R)^2 + dyy_1(Y/R)^2 \\ &\quad + dzz_1(Z/R)^2 - 3dxy_1(X/R)(Y/R)^2 \\ &\quad + dxz_2(X/R)(Z/R)^2 - dyx_2(Y/R)(X/R)^2 \\ &\quad + dyz_2(Y/R)(Z/R)^2 + dzx_2(Z/R)(X/R)^2 \\ &\quad + dzy_2(Z/R)(Y/R)^2 \\ &\quad - dxb/2(X/R)\text{sign}(dX/dt) \\ &\quad - dyb/2(Y/R)\text{sign}(dY/dt) \\ &\quad - dzb/2(Z/R)\text{sign}(dZ/dt) \\ &\quad - \beta_1(Z/R)(X/R) + \alpha(X/R)(Y/R) \\ &\quad - \beta_2(Y/R)(Z/R) + M_{xy}(X/R)(Y/R)\text{Dir} \\ &\quad + M_{yz}(Y/R)(Z/R)\text{Dir} \\ &\quad + M_{zx}(Z/R)(X/R)\text{Dir}\end{aligned} \quad (13)$$

where the first terms of polynomial functions are considered for each parametric error components for the efficiency of calculation, and Dir is a function returning plus(Clockwise) or minus(Counterclockwise) depending on the rotation of kinematic ball bar. For the volumetric error equations, the roll error components and the effect of tool offsets are not considered here for simplicity of modeling.

4.3 LEAST SQUARES TECHNIQUE FOR THE ERROR DIAGNOSIS

When the ball bar measurement data, ΔR , are provided, there are several error components involved as in eqs(13). In order to analyse the error components from the ball bar measurement data, the least squares technique has been effectively applied in this paper.

When ΔR_m is the circular error to be modeled, eqs(13) can be expressed as the linear combination of polynomial functions.

$$\Delta R_m = \sum A_i F_i \quad (14)$$

where F_i is the i th polynomial function, and A_i is the coefficient of the i th polynomial function.

Let E as the sum of squares of deviation between ΔR and ΔR_m , then

$$E = \sum (\Delta R - \Delta R_m)^2 \quad (15)$$

Applying the variational principle to eq(15) for minimizing E ,

$$\partial E / \partial A_i = 2 \sum (\Delta R_m - \Delta R) \partial(\Delta R_m) / \partial A_i = 0 \quad (16)$$

From eq(14), $\partial(\Delta R_m) / \partial A_i = F_i$, thus eq(16) becomes

$$\sum \Delta R_m F_i = \sum \Delta R F_i \quad (17)$$

If we set $f = [F_1, F_2, \dots, F_n]^T$, $a = [A_1, A_2, \dots, A_n]^T$, then eq(17) becomes

$$(f^T f) a = (\Delta R f) \quad (18)$$

Eq(18) is a typical linear equation for unknown matrix, and therefore it can be solved by applying numerical methods such as Gauss elimination method, etc.

5. PRACTICAL APPLICATION

The developed error analysis technique has been practically applied to a column type machining center, which was installed on the shop floor in the Seoul National University.

A commercially available kinematic ball bar of 150mm nominal length has been used for the circular error measurement, and Fig.2 shows a typical setup for the ball bar measurement. The ball bar has been precalibrated with respect to a calibration fixture made of Zerodour, prior to performing measurement. After the measurement path has been planned, the appropriate CNC code has been generated for the circular contouring, then it has been downloaded on the CNC machine tool. The measured data are then analysed and the error diagnosis has been performed.

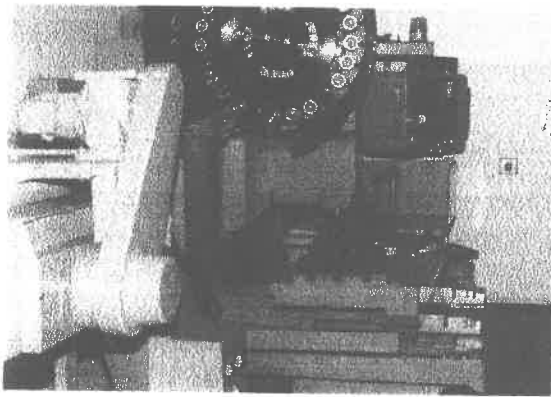


Fig.2 SETUP FOR THE BALL BAR MEASUREMENT

Fig.3 shows the raw data plot for the circular error measurement in XY, YX, and ZX planes, respectively, where the dotted line indicates the counter clockwise contour and the solid line does the clockwise contour. The raw data were analysed by the developed system and the error components were evaluated. The analysed error components were removed from the raw data of the circular measurement, and the residual circular errors were calculated then plotted. Fig.4 shows the residual circular errors after removing the error components, giving remarkably reduced error pattern.

Therefore the developed error analysis system was found as an efficient tool for the error diagnosis of machine tools based on the kinematic ball bar measurement.

6. CONCLUSION

(1) A computer aided analysis system has been developed for the parametric error components, based on the circular error measurement using the kinematic ball bar.

(2) In this paper, a new approach has been proposed and tested such that the 3D volumetric error model is effectively integrated for the ball bar measurement, thus the 3D volumetric errors has been efficiently

assessed with the ball bar measurement data.

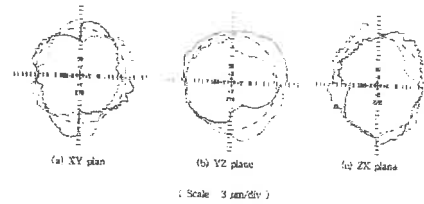


Fig.3 RAW DATA PLOT FOR THE BALL-BAR MEASUREMENT
(COLUMN TYPE MACHINING CENTER)

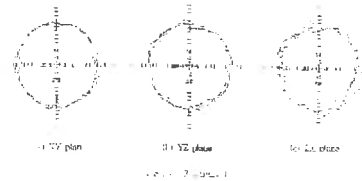


Fig.4 RESIDUAL ERROR PLOT AFTER ANALYSING
ERROR COMPONENTS
(COLUMN TYPE MACHINING CENTER)

REFERENCE

1. Bryan, J., A simple method for testing measuring machines and machine tools, Precision Engineering, Vol.4(2), 1982
2. Knapp, W., Test of three dimensional uncertainty of machine tools and measuring machines and its relation to machine errors, Annals of CIRP, Vol.32(1), 459-464, 1983
3. Kakino, Y., et al, The measurement of motion errors of NC machine tools and diagnosis of their origins by using telecopying magnetic ball bar method, Annals of CIRP, Vol.36(1), 377-380, 1987
4. Burdekin, M., Park, J. A computer aided system for assessing the contouring accuracy of NC machine tools, Proceeding of the 28th MATADOR conference, Manchester, 1988
5. Kunzmann, H., et al. On testing coordinate measuring machines with kinematic reference standards, Annals of CIRP, Vol.32(1), 465-468, 1983
6. Park, H., Development of computer software for volumetric error map integrating the machine geometric error and the probe error in pursuit of a numerical virtual CMM, J. of KSME, Vol.20(3), 945-959, 1996 (in Korean)
7. Y. Koren, Computer Control of Manufacturing Systems, McGraw-Hill Co., 1983
8. Renishaw Ballbar Diagnostic Manual, Renishaw Co. 1992

ACKNOWLEDGEMENT

The authors are grateful to the NSDM-ERC of Pusan National University for the partial funding to the research.

Analysis of Tool Trajectories (2D Tool Paths) in Machine Tools

Prof. Dr. F. Rehsteiner, S. Weikert
Institute of Machine Tools and Manufacturing (IWF)
Swiss Federal Institute of Technology (ETH)
Zürich, Switzerland

Abstract

At the Institute of Machine Tools and Manufacturing (IWF) methods have been developed to investigate the performance of machine tools particularly with regard to high feed rates by means of the newly introduced cross grid system distributed by Heidenhain.

The measuring grid is well suited for observing the relative motion of two bodies in planar motion within a circular range of 160 mm. It is based on the same measuring principle as the well-known glass scales, and thus it features the same high resolution and accuracy. In its application to machine tools the two components, i.e., the grid plate and the optical scanning head, are fixed to the tool holder (in case of milling machines: the spindle head) and to the workpiece table, respectively. This enables them to observe geometric and dynamic effects of machine tool performance on the tool path. In the present paper a number of such effects are analyzed.

Keywords:

Dynamic Measurements, Machine Tool Metrology, Measuring Grid, Axis Cross-Talk

Introduction

With the growing use of high speed cutting methods the importance of the dynamic performance of machine tools is increasing steadily. While at low feed rates the tool path errors are caused primarily by geometric inaccuracies of the machine tool due to inaccuracies in part geometry, assembly, and thermal effects, at high feed rates structural deformations due to high accelerations and path deviations due to inadequate control and feed drive performance become increasingly important.

This new situation calls for new, adequate methods for monitoring the tool path in real time in order to allow a complete assessment of machine tool performance. In the past the most widely used method to do this was to machine suitable parts and use the results of measuring their geometry, dimension, and surface quality to interpret machine tool performance and to improve it. This method is not only time consuming but the results are also influenced by the machining process which tends to introduce additional errors. The method based on the cross grid system, however, does away with these problems and permits a much faster improvement cycle by making the results of modifications to the machine tool and to the controls observable in real time and unobstructed by secondary effects. Additionally, static features like the true straightness of „straight“ paths, angle deviations of rectangular movements, or positioning errors, can also be investigated. The combined information about machine parameters, motion parameters (like, e.g., feed rate and acceleration) and their result in terms of position coordinates as a function of time allow quite a complete assessment of systems performance.

Measuring system

At the heart of the system there is the grid plate measuring 160 mm in diameter which carries two line grids orthogonal to each other, see Fig. 1. A scanning head which is also equipped with an orthogonal line grid simultaneously lights the grid and observes the reflected light. The intensity of the latter varies sinusoidally as the two grids move relative to each other.

The signal picked up is interpolated in 1024 intervals by the measuring circuit which, together with the grid constant of $8\text{ }\mu\text{m}$, results in a resolution of about 5 nm . A personal computer (PC) is then used for storing and further processing the recorded data. Peak monitoring frequency in the system used at the IWF laboratories is 8 kHz allowing 200'000 points (x-y coordinates) to be picked up in a single measurement. In the standard version (ACCOM®-software) up to 8000 points can be stored and processed further by a number of suitable algorithms (ISO 230-4 standard)[1].

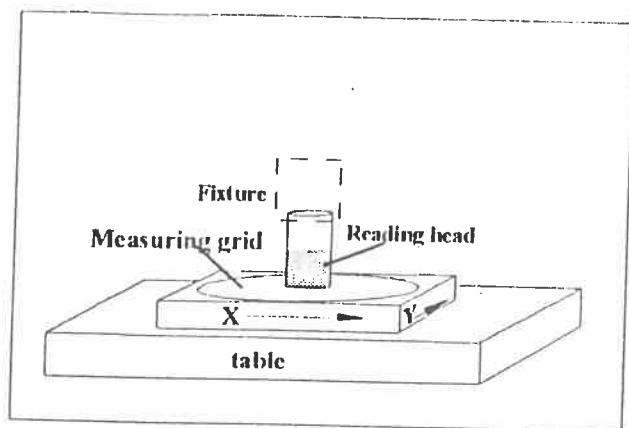


Fig.1: Components of Measuring System

To perform a measurement, first the measuring grid is placed on the machine table with grid lines arranged parallel to the guideways. Second, the scanning head is mounted to the tool fixture at a vertical distance of 0.5 mm from the grid. If the head is mounted in the tool spindle then care must be taken to clamp the latter such that the head can not change its orientation. The angle between the head and the grid is fine-tuned by observing the signal output, whereupon the actual measurements may begin. This whole procedure typically takes about 10 to 15 minutes.

Multiple Measuring Options

A variety of measurements can now be made with the system described above, see Fig. 2.

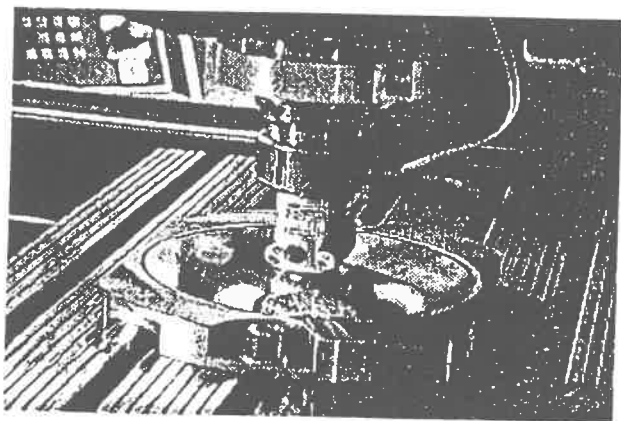


Fig. 2: The measuring system mounted to a machine tool (measuring grid on machine table, measuring head in tool fixture)

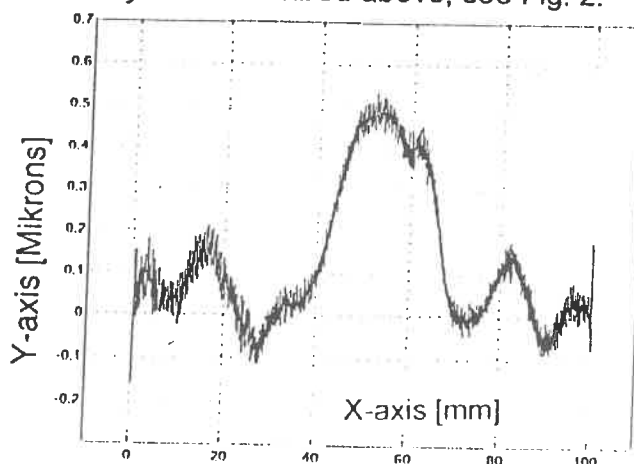


Fig. 3: Measured path for a straight line parallel to the X-axis at a feed rate of 4 m/min .

Straightness

A very simple test path, i.e., a straight line parallel to one of the machine axes which is to be run at a given constant feed rate is shown in Fig. 3. The high resolution and fast read-out of the measuring system not only allow to monitor path deviations in the direction of feed as could be done also by laser interferometry, but also transverse ones. These are direct indications about the quality of the mechanical machine elements like the straightness of the feed spindle and of the guideways.

Cross-talk in reversing motion

Moving along a path in one direction and then reversing it will yield deviations in the axial as well as the lateral direction that are specific to this reversal. Results of such a test are shown in Fig. 4: Movement starts at point 1 in the positive direction; at point 2 it stops for 2 seconds. Then it goes on to position 3 and, without stopping, back to position 2 where there is again a stop of 1 second. Next, point 1 is again approached in the original (positive) direction and a stop of 1 second is put in. Finally it moves to point 1 via point 5 (no stop) and a new cycle starts. The lateral displacement between paths obtained by movement in opposite directions is clearly visible.

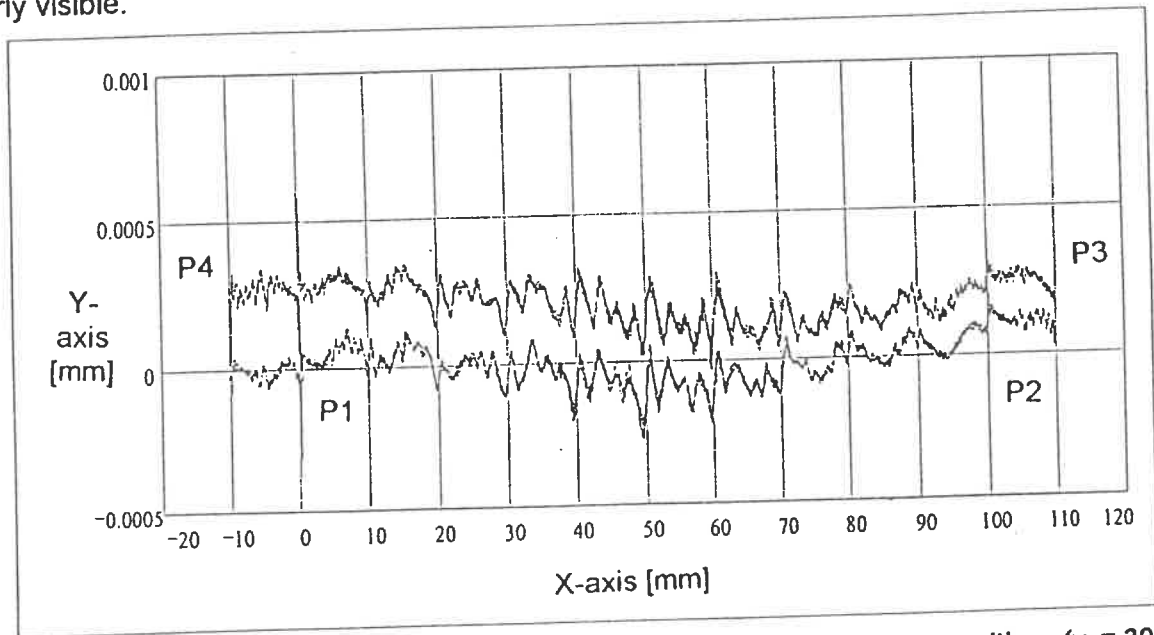


Fig. 4: Back-and-forth motion along a straight axis with short stops at various positions ($v_f = 3000$ mm/min)

In addition there is a strong lateral deviation occurring at the rate of spindle pitch which, here, is 10 mm. Contrary to *axial* deviations caused by spindle pitch errors which can be detected easily by laser interferometry and which do not show up in the machined surface this error enters surface quality full size.

Displacement due to thermal effects

If the scheme described in the previous section is continued during a prolonged time there is an additional displacement of the whole motion due to non-uniform heating of the structural elements. This is shown in Fig. 5 for the y direction. Displacement reaches a maximum after about one hour but gets reduced somewhat later on. However, the short term displacements remain practically identical during the whole experiment. By observing the displacement occurring even if no axis movement takes place one can observe also the thermal effect of temperature variations in the surroundings of the machine.

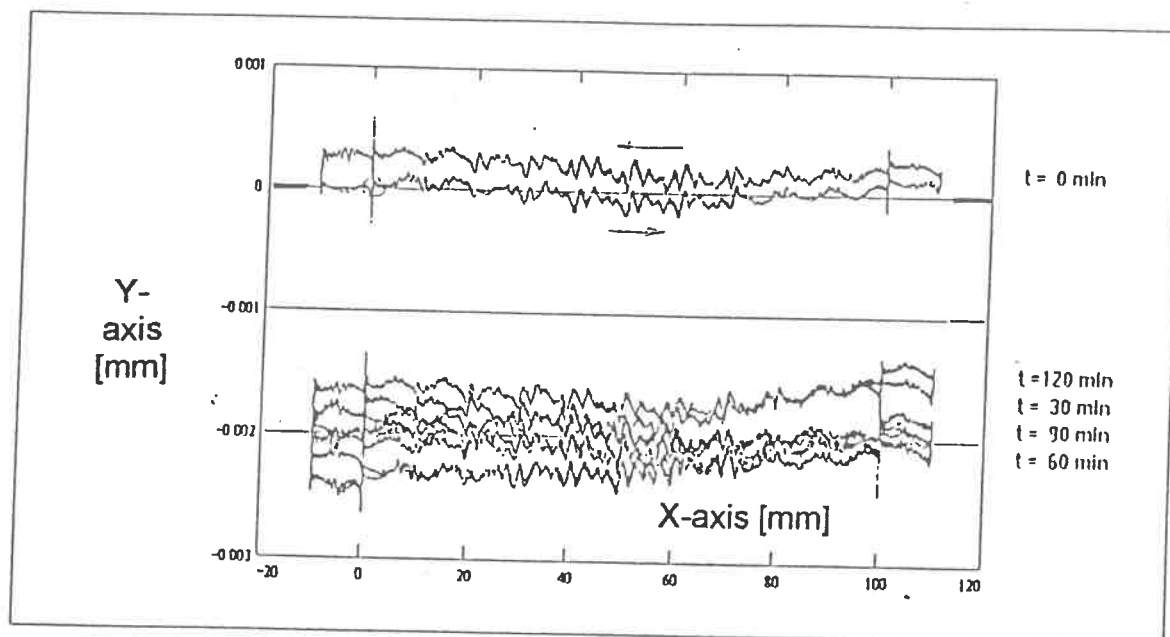


Fig.5: Back-and-forth motion along a straight axis with longer stops at various positions ($v_f = 3000$ mm/min) over two hours (shown graphs measured at the start, after 30, 60 90 and 120 minutes)

Another application referring directly to machine dynamics is „cornering“ (moving along a path with a pointed, rectangular corner) where the amount of overrun at various parameter settings is of particular interest. This is shown in Fig. 6 for running around a sharp corner at a programmed speed of 6000 mm/min. Some overrun is clearly visible. The shape of the path can also be interpreted in terms of speed and acceleration and thus yields a lot of information about the quality of machine tool dynamics.

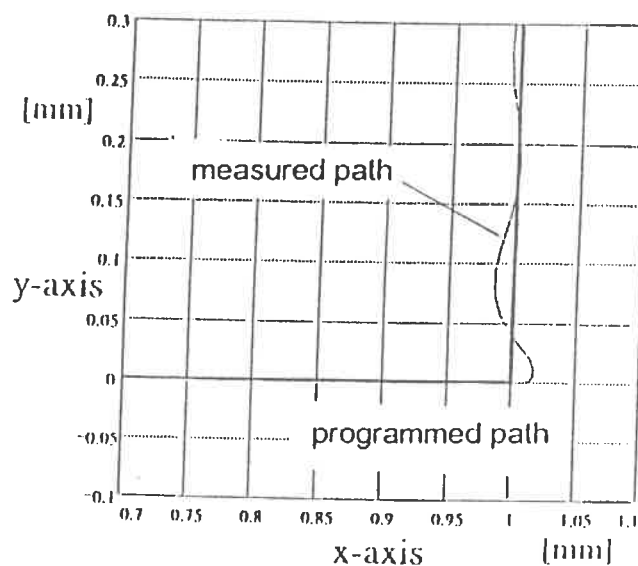
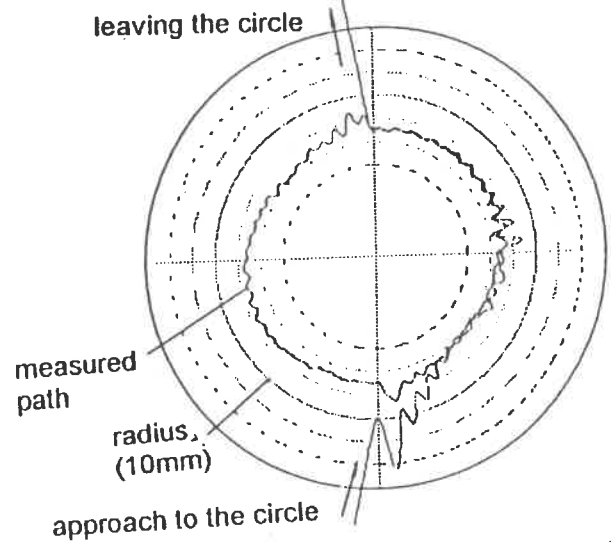
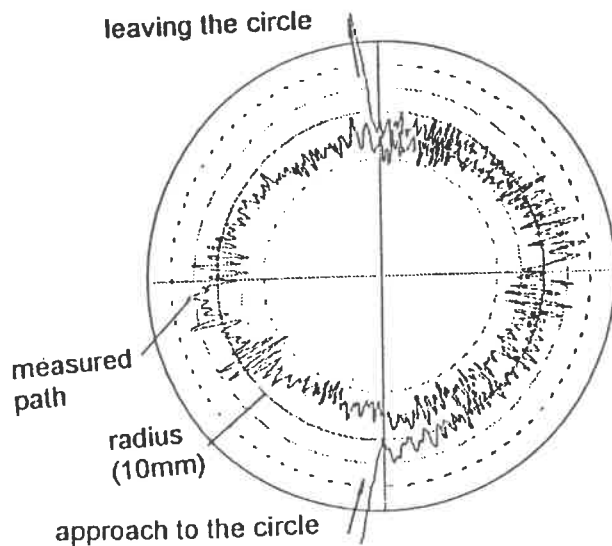


Fig. 6: Cornering an edge at $v_f = 6000$ mm/min.

Circular test

In the classical „circular test“ (using a ball bar or tracing a master) a circular test piece of high quality is traced by a mechanical tracer and the deviation of the recorded path from a true circle is then used to interpret machine quality [2]. Tracer friction respectively inertia limit this method to fairly low speeds. In contrast the optical tracing method permits much higher tracing speeds; the system used here could be used up to a feed rate of $v_f = 24$ m/min with 5 nm resolution. Figs. 7 and 8 show results obtained on two different machines for identical motion inputs (circle with 10 mm diameter to be followed at $v_f = 6$ m/min). Fig. 7 was obtained on a highly dynamic machine: Good average accuracy but strong radial oscillations, whereas Fig 8 shows corresponding results for a less dynamic but „good-natured“ machine. There is a substantial error all along the path but little oscillations.



Figs 7, 8 Results of measurements performed at a feed rate of 6 m/min on two different machine tools. The theoretical path is a circle of 10 mm radius. Line separation is 10 μ m.

Axis cross talk

By the displacement of machine parts relative to each other due to machine tool motion its mass distribution changes constantly. Thus, accelerations create different, configuration dependent displacements which may point in directions other than the feed direction (cross talk). These are, however, often poorly recognizable by the standard (linear) measuring systems on machine tools since they occur outside the measuring/control loop.

Fig. 9 shows the path that was observed on a large machining center the axis configuration of which is shown in Fig. 10 when a straight line parallel to one machine axis was to be run. Lateral displacements proportional to the acceleration in the feed direction can be seen clearly. The reason is the large distance between the point of feed force input at the feed spindle and the center of gravity of the moving column. As the lower curve shows the programmed peak acceleration of 2 m/sec² in the z direction was realized correctly.

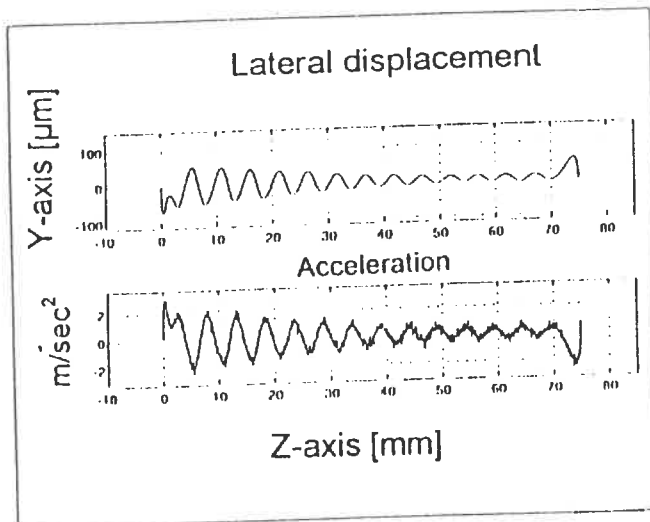


Fig.9: Acceleration and path-deviation along a straight line 100 mm long. Top: x-y plot of measured data, Bottom: acceleration over z-axis

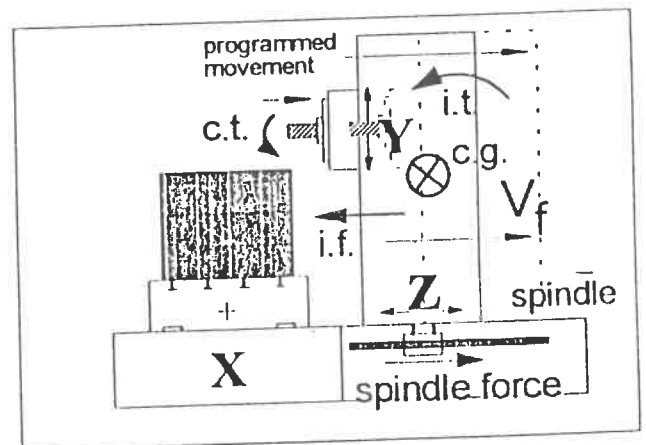


Fig.10: Axis configuration of a large machining center and programmed path for the measurements of Fig. 9 (c.g.=center of gravity; i.f.=inertia force; i.t.=inertia torque; c.t.= crosstalk motion of tool)

Comparison of results by the cross grid system and surface measurements

Consider a surface created by face milling. A comparison of the surface contour as measured by conventional methods (stylus tracing) and the one obtained from grid measurements by incorporating cutter geometry shows good agreement between the two. Such a comparison is shown in Fig. 11. Features with a long wave length agree well while those with a short wave length differ substantially. Apparently the reason for this is the poor roundness of the two flute cutter used in this test. Although the cutting conditions were quite conventional (cutter diameter 10 mm, feed 5 mm, cutting speed 720 m/min) their influence on the resulting part geometry was small. However, both measuring methods show well the influence of spindle pitch, see Fig. 11.

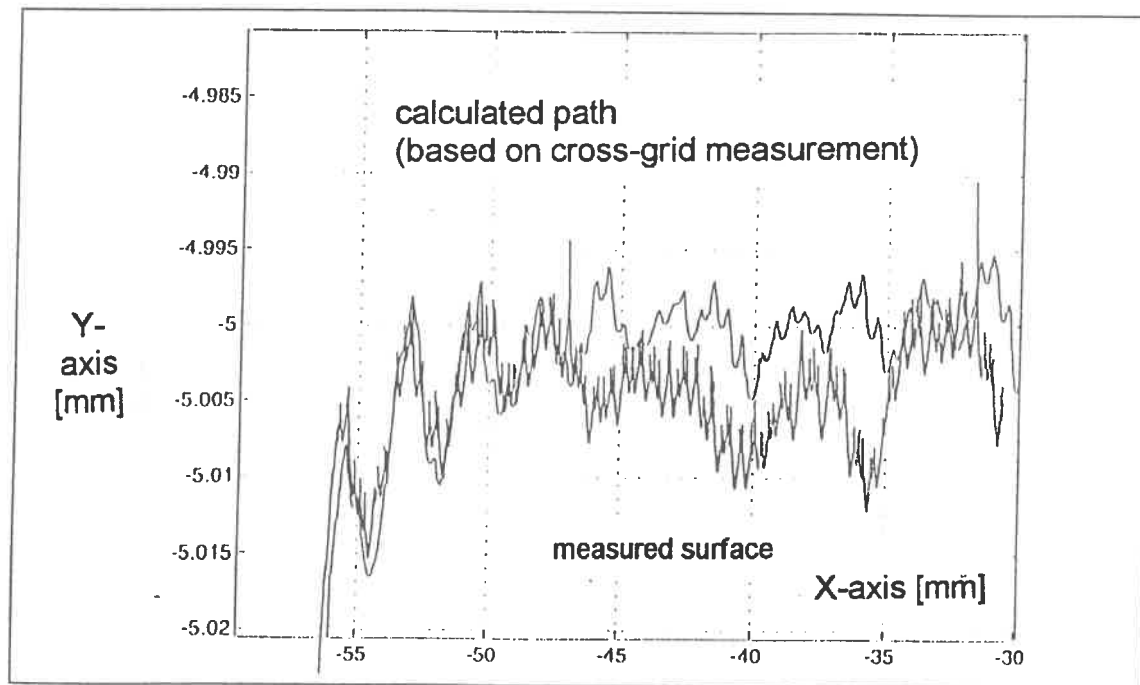


Fig. 11: Comparison of a measured surface, all the theoretical contour, that was calculated, based on a cross-grid measurement

Conclusion

As a new tool for investigating machine tools both with respect to static and dynamic performance the cross grid system offers a large number of possibilities.

Particularly in high speed machining where the cutting force is small the actual surface contour of the workpiece can be deduced to a high degree of validity from grid measurements. Thus, the actual tool path which has been influenced by the controls and control algorithms, by the drive and structural characteristics of the machine tool can be observed readily without resorting to cutting tests. This reduces strongly the effort needed to optimize machine performance - e.g., by adapting control parameter settings.

References

- [1] HEIDENHAIN-Information, „Messgeräte zur Abnahme und Kontrolle von Werkzeugmaschinen“ 5/95
- [2] Knapp W., Hrovat S. Der Kreisformtest zur Prüfung von NC-Werkzeugmaschinen Eigenverlag S. Hrovat, 1986
- [3] ISO 230-2: Acceptance code for machine-tools - Part 2: Determination of accuracy and repeatability of positioning of numerically controlled machine tools. International Organisation of Standardisation, 1988

Comparison of Control Strategies for Piezoelectric and Magnetostrictive Micropositioning Devices for Ultra-Precision Machining

J. C. Campos Rubio¹, J. G. Duduch¹, A. E. Gee²,
A. J. Vieira Porto¹ and H. Lêdo¹

¹ Escola de Engenharia de São Carlos (EESC) - Universidade de São Paulo (USP)
Departamento de Engenharia Mecânica (SEM) - LAMAFE / Ultra-Precisão
P.O. Box 359 CEP 13560-254 - São Carlos - SP - Brazil

² School of Industrial and Manufacturing Science
Cranfield University - Cranfield, Bedfordshire MK43 0AL - England

Abstract

Ultra-precision machining to sub-micrometre depths is shown to be a threshold area of work requiring a combination of the best of materials, sensors, positioning devices (actuators) and control strategies. In ultra-precision design, it is extremely likely that there will be few design options for selection. In the case of ultra fine tool feed, the piezoelectric actuator is a potential choice. Currently, European and Japanese tool manufacturers are investigating magnetostrictive actuators for this purpose. Both piezoelectric and magnetostrictive materials suffer from hysteresis type non-linearity, so that the output of such systems depends upon the previous input, and absolute positioning is only achievable with the aid of feedback control. This work compares some modern control techniques for the positioning of a tool post for ultra-precision machining of brittle materials (in the nanometric range), eg, lead-lag Filter, PID with feedforward and fuzzy logic. The performance of the micropositioning device (tool post) using both piezoelectric and magnetostrictive actuators is assessed by means of simulation techniques. The predicted performance is compared with results obtained by other authors.

Key words: Micropositioning Devices, Piezoelectric Actuators, Magnetostrictive Actuator, Control Strategies.

1. Introduction

In recent years, ultra precision machining to submicrometre range has gained increasing importance. The fabrication of components with high form accuracy (of the order of 0.1 μm) and surface roughness of the order of 10 nm Ra (eg scanner mirrors, infrared optics, lens molds, computer memory disks etc) has demanded the use of machines with higher positioning precision of the order of nanometres (Dow et al, 1991).

Currently, efforts have been made to machine ceramics and vitreous materials using single point diamond tools where the material is removed under certain restricted conditions which permit plastic removal, resulting in surfaces free of sub-surface damage. This ductile mode regime is obtained by controlled depths of cut and feedrates.

Together with the necessity of controlling the parameters involved, there is also the need to correct systematic errors and to attenuate disturbances in the cutting process. This can be done by correcting the movement of the cutting edge through real time control of the position of the tool via a position feedback signal from a sensor which controls directly a tool holder, before or during the process.

It is therefore necessary to use a servo-controlled micropositioning system (micro tool servo drive) which permits low amplitude movements and low following error to high frequencies. One of the basic components of these systems is the actuator which changes an electrical signal into movement and is based on solid state technology.

To reach the requirements quoted above, the next generation of machine tools will have to possess better accuracy and repeatability characteristics added

to adequate dynamic performance to make possible the economic production of highly complex components with superior metrological characteristics and made of several difficult-to-machine materials.

This work proposes the use of modular micropositioning systems which can be incorporated to existent machines as well as to new designs. Two types of solid state actuators will be considered: piezoelectric and magnetostrictive. The dynamic performance of these actuators will be compared using different closed loop digital control strategies through numerical simulation.

2. Micropositioning device and prototype

Most commercial precision machines do not have repeatability and accuracy capabilities to machine brittle materials. The development of positioning/correction systems of the tool is important as it is ultimately responsible for the form, finish and sub-surface integrity of a workpiece.

Several devices have been proposed and developed. Figure 1 shows a modular type of tool holder which uses a flexure support system. (Kohno et al, 1989)

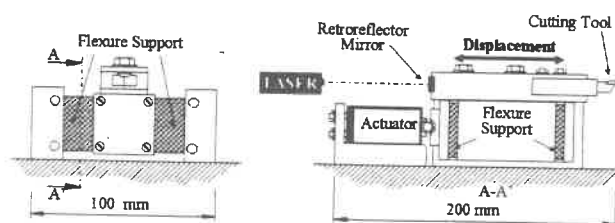


Fig. 1 - Modular Toolholder for Micropositioning

3. Solid State Actuators

Solid state actuators are a special group of linear positioners whose principle is based on strain phenomena. This type of actuator is recommended for high precision positioning. Other types of linear actuator could be presented in this work (eg, eletromagnetic, hydraulic, pneumatic), but they are hardly used to solve problems of positioning in the submicrometre range (Campos et al, 1996 a).

Piezoelectric and magnetostrictive actuators are presented below.

3.1 Piezoelectric Actuators

Piezoelectric actuators are based upon the properties that certain crystalline materials possess to expand or contract when a voltage is applied to them. The

opposite mode exists: when a mechanical pressure is applied to the crystal a voltage is produced on its surface. Crystalline materials normally used are sinterized ceramics such as barium titanate, lead titanate-lead zirconate (PZT).

One common type of piezoelectric actuator is the PZT stack. In this arrangement, a number of plates and electrodes are stacked together. The number of plates determines the stiffness and stability of the actuator.

The advantages of piezoelectric actuators are:

- Nanometric positioning resolutions
- No backlash
- High efficiency
- High load carrying capacity
- Large bandwidth (limited by the sound speed in the crystalline material (Physik Instrumente, 1992)

Piezoelectric actuators, however, do not support loads out of centre and cannot be loaded in tension (some are pre-loaded). They are sensitive to temperature variations and present up to about 15% hysteresis and non-linearities. Feedback control is used to overcome the problems of non-linearities and hysteresis (Slocum, 1992).

3.2 Magnetostrictive Actuators

Originally developed as part of a US Navy project to find new sonar materials, magnetostrictive materials transduce low voltage electrical energy into mechanical movement (Edge Technologies, 1995).

The term Magnetostriction refers to the physical phenomenon of dimensional variation that occurs in a magnetic material when a magnetic field is applied (Guillemin effect).

Magnetostrictive translator (MST) are built from an alloys of metals, terbium, dysprosium and iron.

Among the characteristics that make this kind of actuator one of the best solutions, one can mention:

- High repeatability
- Reliability
- High load carrying capacity
- Large bandwidth
- scarcely affected by environmental changes

Its high performance makes this type of actuator an excellent alternative for micropositioners, vibration control, seismic sources, noise control. However it suffers from thermal drift.

Both MST and PZT materials work in a way to transform an electric signal into linear motion in the

range of a few micrometres. The load carrying capacity can easily reach the order of kN and frequencies can range from zero to several kHz. Due to their excellent positioning precision and speed stability, these actuators can be utilized in the positioning of ultra-precision machine tools (Schäfer and Janocha, 1995).

4. Modeling of Micropositioning Device

Each element of the dynamic system will be described by its respective mathematical model. The block diagram of Figure 2 shows the structure of the dynamic system.

This system may be divided into two parts: the digital controller and the plant which comprizes the other elements (eletromechanical components).

On the plant, two parts can be clearly distinguished, one is the actuator (PZT or MST) and the other is the moving mechanical elements.

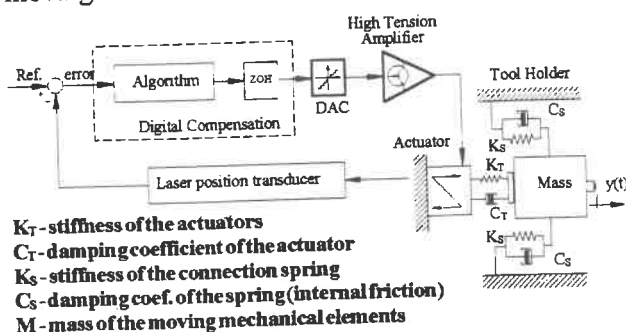


Fig. 2 - Structures of Dynamic System

Mechanical Elements (Tool holder)

The mathematical model of the moving mechanical elements of the tool holder is shown in Equation 1 where ζ is the damping ratio (0.08) and ω is the resonant frequency of the system (200 Hz) for a mass of 650 g. (Miron, 1989 / Sato et al, 1994).

$$\frac{Y(s)}{X(s)} = \frac{2 \cdot \zeta \cdot \omega_c \cdot (s + \frac{\omega_c}{2 \cdot \zeta})}{s^2 + 2 \cdot \zeta \cdot s + \omega_c^2} \quad [01]$$

Actuators

These devices convert an electrical input into a displacement. In this model, it is assumed that the actuator displacement is proportional to voltage/current input. The actuator may, therefore, be described as a second-order system where ζ is the damping ratio and ω_n is the resonant frequency (ω_n

$= 2\pi f_r$) (Sanchez et al, 1992 / Okazaki, 1988 / Tanaka et al, 1994).

- **PZT** - In this work is considered a Physik Instrumente P-239-10 piezoelectric actuator, with 10 μm expansion at -1000 volts.

$$\frac{X(s)}{V(s)} = \frac{K_{PZT} \cdot \omega_n^2}{s^2 + 2 \cdot \zeta \cdot \omega_n + \omega_n^2} \quad [02]$$

- **MST** - An MP 50/6 Etrema Terfenol-D^R magnetostrictive actuator is considered with +25 μm expansion at + 1.5 ampere.

$$\frac{X(s)}{I(s)} = \frac{K_{MST} \cdot \omega_n^2}{s^2 + 2 \cdot \zeta \cdot \omega_n + \omega_n^2} \quad [03]$$

Both piezoelectric and magnetostrictive materials suffer from hysteresis type non-linearity, so that the output of such systems depends upon the previous input, and absolute positioning is only achievable with the aid of feedback control. There appear to be two regimes of operation, one for several micrometre range where the effect is strictly 'amplitude dependent' and that below this level where the effect diminishes with reducing amplitude. The former can be modelled as a (fixed) attenuation and phase lag; the latter indicates that for ultra-precision control systems this effect is negligible (Duduch and Gee, 1990)

Table 1 - System parameters

	Tool holder	PZT	MST
freq. resp. (f_r)	450 Hz	10 kHz	5 kHz
stiffness (k)	$5.25 \cdot 10^6$ N/m	$140 \cdot 10^6$ N/m	$7 \cdot 10^6$ N/m
damping ratio	0.080	0.700	0.097
gain factor(K_{iii})	452.39	$7.66 \cdot 10^{-7}$ m/V	$1.66 \cdot 10^{-5}$ m/A

5. Measurement System

As shown earlier, solid state actuators have limited repeitibility due to non-linearities which makes necessary the use of position feedback control to improve performance.

Comercial micropositioning devices normally have a built-in position sensor such as capacitance micrometer or strain gauges.

In this work, a Laser Diode (GaAlAs) Michelson Interferometer was chosen. This choice gives high resolution (1.5 nm, $\lambda/512$ with $\lambda = 780$ nm) and

complies with Abbé principle as well as permitting mounting on an independent metrological frame (Gee et al. 1989).

6. DSP-Based Board Controller

The board used for the digital control of movement is based upon DSP TMS320C26 of Texas Instruments. It provides easy communication with the host computer (PC-AT 386), permitting real time analyses of system performance, digital/analog input/output with 14 bits resolution and direct connection with power amplifiers.

It is possible to implement various compensation algorithm.

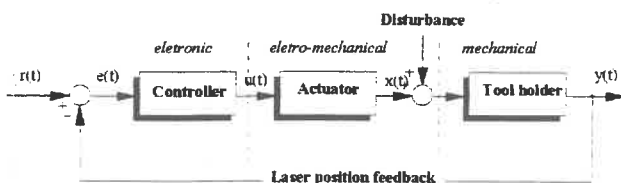


Fig. 3 - Block diagram of closed-loop control system

7. Displacement Control System

The next step is to propose a closed-loop position control system adequate for the system described above and able to meet all (ultra-precision) required specification. Lead-Lag Filter, PID, PID - feedforward and Fuzzy Logic will be considered.

• .Lead-Lag Filter

A direct digital controller will first be considered, based upon the knowledge of its analog equivalent (Lead-Lag filter).

• .PID

In the PID filter, the proportional term contributes to improve the frequency response, the derivative term improves stability and the integral term diminishes the following error

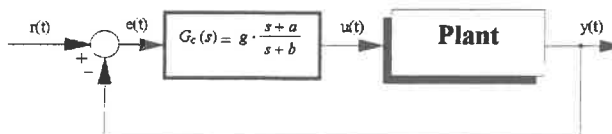
• .PID - feedforward

The integral term of a PID control system reduces the following error on the one hand, but decreases stability on the other. Using a "feedforward" loop, it is possible to reduce steady-state error without, however, reducing stability (damping).

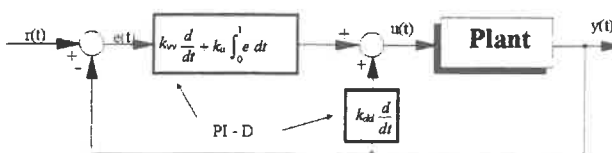
• .Fuzzy Logic

A Fuzzy Logic-based controller is proposed in this study. A Fuzzy Logic controller basically has the following characteristics: it uses linguistic variables instead of or in addition to numerical ones; it is determined by simple relationships between variables establishing "fuzzy" condition; and complex relationships between variables are represented by "fuzzy" algorithms.

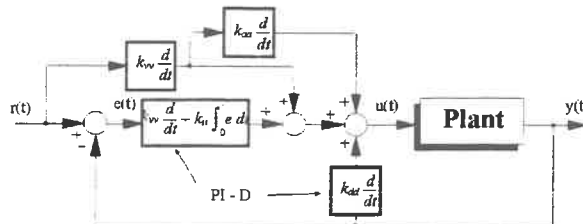
This fuzzy controller is based on classical models where the main part is a "look up" table, previously determined. This table corresponds to a matrix where the inputs are position errors and position-error variation. Thus, pair of input elements corresponds to an element of the matrix, whose value gives a quantitative measure of how is the system compared with the behaviour desired (Li and Lau, 1989).



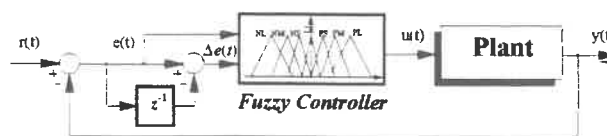
(a) Lead-Lag Filter



(b) PID



(c) PID - feedforward



(d) Fuzzy Logic

Fig. 4 - Digital Filters

8. Numerical Simulation

Three parameters will be used to assess the servo-controllers' behaviour: (a) transient response (as quick as possible, without overshoot), (b) rejection

ability of disturbance to attenuate noises occurring (e.g. due to cutting process), and (c) dimensional accuracy of complex contours - ability to follow variable (with time) contours with little or no error. The model proposed (in section 4) is a preliminary approximation and was obtained analytically. There are no experimental results as yet. For this reason, control techniques were chosen to satisfy the control requirements set out above, without taking into consideration optimization, self tuning, etc. The parameters of the digital filters were selected through a sensitivity study of the loop and the sampling time value was set as 300 μ sec. for all cases. The effect of the digital/analog conversion was considered as it limits the limits the resolution. Figures 4, 5 and 6 show the closed-loop response obtained for different types of excitation. The graphs were plotted through simulation of the dynamic system and controller via MATLAB^R simulation software.

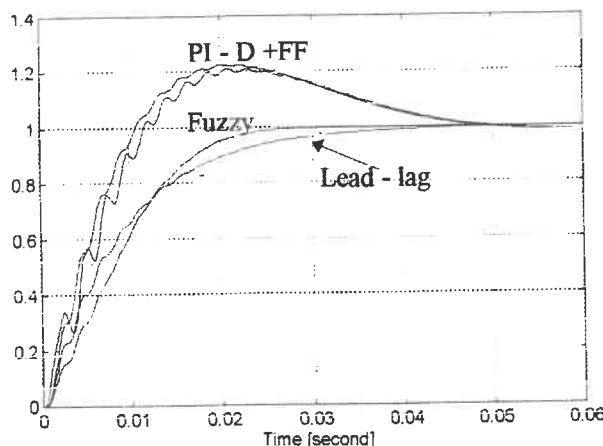


Fig. 4. - Transient response (unit step)

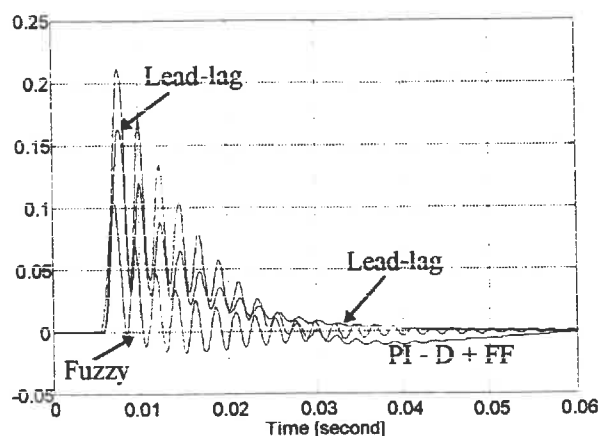


Fig. 5. - Disturbance rejection ability

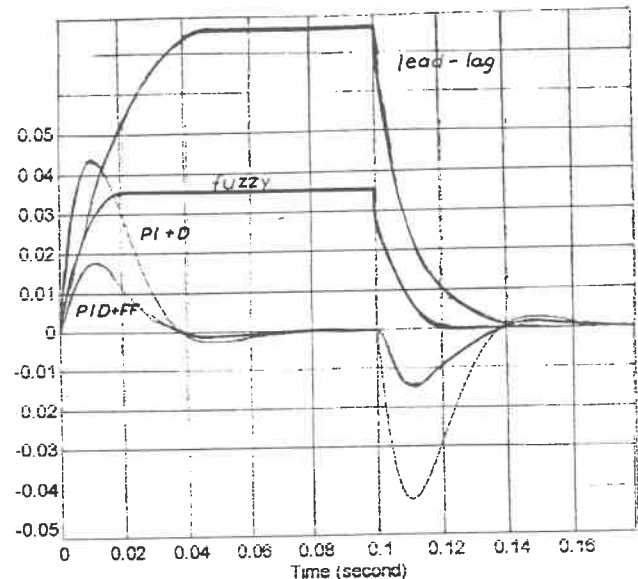


Fig. 6. - Trajectory tracking

9. Summary and Conclusions

This work is intended to show the development of a micropositioning system to be used in ultra-precision single-point diamond turning. The system utilizes the micro-displacement characteristics of solid state actuators. Generally, it can be said that this type of actuator greatly facilitates the design of a micro-tool holder. The mathematical model proposed, although neglecting non-linearities such as static friction, variable stiffness and damping, represents fairly well the dynamic behaviour of the electro-mechanical dynamic system.

The dynamic system obtained permits changing the actuator to suit the desired characteristics and needs, making the system modular and easily re-configurable.

The comparison of the various control strategies shown here allows us a better insight of the system behaviour and judgement of which modifications can be performed for the correct performance.

Finally, the results obtained in this study point to the possibility of knowledge-based techniques, such as Fuzzy Logic and Neural Networks to control magnetostrictive actuators, as these techniques do not require detailed mathematical models to formulate the algorithms, and have more adaptive capability to the variations intrinsic to the process.

10. References

- Campos Rubio, J. C., Duduch, J. and Porto, A.;** Acionamento Direto de Ferramentas de Corte em Máquinas de Ultra-Precisão. *Anales CIM, Valdivia (Chile)*, pp 220-223, 1996.
- Dow, T., Miller, M. and Falter, P.;** Application of Fast Tool Servo for Diamond Turning of Nonrotationally Symetric Surfaces, *Precision Engineering*, Vol. 13, N°4, pp 243-250, 1991.
- Duduch, G. and Gee, A.;** *Proc. ASPE*, Vol. 2, Rochester, NY, pp 170-173, 1990.
- Edge Technologies, Inc.,** Etrema Terfenol-D Magnetostrictive Actuators (catalog), Ames, IA, 1995.
- Gee, A. McCandlish, S. and Puttick, K.;** On-axis Interferometric Measurement for Tool Positioning in Ultra-Precision Single-Point Machining. *Proc. SPIE*, Vol. 1121, pp 74-80, 1989.
- Kohno, T., Okazaki, Y., Ozawa, N., Mitui, K. and Omoda, M.;** In-Process Measurement and a Work-Piece -Referred Form Accuracy Control System (WORFAC): concept of the method and preliminary experiment, *Precision Engineering*, Vol. 11, pp 9-14, 1989.
- Li, Y. and Lau, C.;** Development of Fuzzy Algorithms for Servo Systems, *IEEE Control Systems Magazine*, April, pp 65-71, 1989.
- Miron, D.,** Design of Feedback Control Systems, HBS Publishers, 1989.
- Okazaki, Y.;** A Micro Positioning Tool Post a Piezoelectric Actuator for Diamond Turning Machines (in Japanese), *JSPE*, Vol. 54, N°7, pp 163-168, 1988.
- Physik Instrumente;** Products for Micropositioning (cataloge) edition E, 1992.
- Sanchez, L.J. Albertazzi, A. and Barczak, C.;** Sistema de Ativo de Estabilização Aplicado a Holografia Eletrônica, *Anales CIM, Santiago (Chile)*, pp 513-518, 1994.
- Sato, K., Obata, F., Matsubara, T. and Mizumoto, A.;** Active Chatter Suppression of Slender Boring Bar Using Piezoelectric Actuators. *JSME International Journal, Series C*, Vol. 38, N°1, pp 36-41, 1995.
- Schafer, J. and Janocha, H.;** Compensation of Hysteresis in Solid-State Actuators, Sensors and Actuators, A 49, pp 97-102, 1995.
- Slocum, A.;** Precision Machine Design. Prentice-Hall, Englewood Cliffs, NJ, 1992.
- Tanaka, H., Murayama, Y., Imada, S. and Shimokhobe, A.;** Control and Elimination of Lead Screw Backlash for Ultra-Precision Positioning. *JSME International Journal, Series C*, Vol. 37, N°3, pp 601-606, 1994.

A Study on Feeding Control Mechanism Analysis of Industrial Lock Stitch Sewing Machine

Keyoung Jin Chun, Dae Young Shin

Korea Academy of Industrial Technology

371-36, Kasan-Dong, Kumchon-Gu, Seoul, Korea

1. Introduction

This paper deals with a kinematic analysis of the feeding control mechanism (FCM) and design variable's effects for rotational angles of the feed regulator link in an industrial lock stitch sewing machine. Relations of the feeding distance of the feed regulator dial shaft and the rotational angle of the feed regulator link are studied. The sensitivity data of each design variable used for tolerances are obtained. The mutual relations of the design variables are discussed from these sensitivity data.

2. Kinematic analysis of the FCM

The FCM has a sliding contact mechanism (SCM) and a four bar linkage (FBL) as shown in Fig. 1.

The origin of a stationary Cartesian X-Y reference frame is on the rotational axis of the feed regulator, which is the absolute origin.

2.1. Kinematic analysis of the SCM

The SCM is shown in Fig. 2. When the feed regulator dial shaft is most close to the absolute origin, D is the distance

between the absolute origin and the center of $C3_u$ or $C3_b$. While the feed regulator dial shaft is moving, d represents the motion of $C3_u$'s or $C3_b$'s center.

The SCM has the two characteristics in the motion as follows. The one is that the feed regulator dial shaft and the feed regulator is contacted by force F and tangent. Here, F has two force directions F_1 or F_2 . The other is that a tangent point is moved by d . While F_1 is being on the feed regulator and d is increasing gradually from 0.0, $C3_u$ contacts with $C1_u$, and then with $L1_u$ after passing arcs. Thus, the kinematic analysis of the SCM could be done at four cases when tangent point is between ① $C3_u$ and $C1_u$, ② $C3_u$ and $L1_u$, ③ $C3_b$ and $C1_b$, ④ $C3_b$ and $L1_b$.

To simplify the analysis of the SCM, two assumptions may be introduced here.

ASSUMPTION 1: Geometries of contacting surfaces of the feed regulator are symmetric.

ASSUMPTION 2: Geometries of contacting surfaces of the feed regulator dial shaft are symmetric.

From ASSUMPTIONS 1 and 2, two

rotational angles (θ_s, θ'_s) of the feed regulator's symmetric axis made by $F_1(\uparrow)$ or $F_2(\downarrow)$ have the same magnitude and the opposite direction each other as shown in Fig. 2. Thus, the kinematic study could be focused on the rotational angle θ_s only while F_1 is acting on.

2.1.1. Analysis of the rotational angle θ_s

2.1.1.1. Kinematic analysis when tangent point is between $C3_u$ and CI_u

When F_1 acts on, $C3_u$ and CI_u are tangent each other as shown in Fig. 3. Two constraint equations are given as

$$f_1 \equiv l_1 \cos(\theta_{cl}) - (r_1 + r_3) \cos(\phi_{cl}) - l_2 \cos(\theta_2) = 0 \quad (1)$$

$$f_2 \equiv l_1 \sin(\theta_{cl}) - (r_1 + r_3) \sin(\phi_{cl}) - l_2 \sin(\theta_2) = 0$$

where $l_2 = \sqrt{(D+d)^2 + d_y^2}$, $\theta_2 = \tan^{-1}\left(\frac{d_y}{D+d}\right)$.

From the equation (1), The equations of two dependent variables θ_{cl} and ϕ_{cl} are derived as

$$\theta_{cl} = \theta_2 + \cos^{-1}\left(\frac{l_1^2 + l_2^2 - (r_1 + r_3)^2}{2l_1 l_2}\right) \quad (2)$$

$$\phi_{cl} = \cos^{-1}\left(\frac{l_1 \cos(\theta_{cl}) - l_2 \cos(\theta_2)}{r_1 + r_3}\right) \quad (3)$$

Thus, θ_s is derived as

$$\theta_s = \theta_{cl} - \eta_1 \quad (4)$$

2.1.1.2. Kinematic analysis when tangent point is between $C3_u$ and LI_u

While F_1 acts on, $C3_u$ and LI_u are tangent as shown in Fig. 4. Two constraint equations are given as

$$\begin{aligned} g_1 &\equiv (r_2 + r_3) \cos(\theta_{c2}) - t_1 \sin(\theta_{c2}) \\ &\quad - l_2 \cos(\theta_2) = 0 \\ g_2 &\equiv (r_2 + r_3) \sin(\theta_{c2}) + t_1 \cos(\theta_{c2}) \\ &\quad - l_2 \sin(\theta_2) = 0 \end{aligned} \quad (5)$$

From the equation (5), The equations of two dependent variables θ_{c2} and t_1 is derived as

$$\theta_{c2} = \theta_2 - \cos^{-1}\left(\frac{r_2 + r_3}{l_2}\right) \quad (6)$$

$$t_1 = \frac{(r_2 + r_3) \cos(\theta_{c2}) - l_2 \cos(\theta_2)}{\sin(\theta_{c2})} \quad (7)$$

Thus, θ_s is derived as

$$\theta_s = \theta_{c2} - \eta_2 \quad (8)$$

2.1.2. Analysis of the feed regulator's the rotational angle θ_i

θ_i is shown in Fig. 2 and derived as

$$\theta_i = \lambda + (-1)^n \cdot \theta_s \quad (9)$$

where $n=2$ when F_1 is acting, $n=1$ when F_2 is acting on, and λ is an angle relative to the symmetric axis.

2.2. Kinematic analysis of the FBL

Fig. 5 shows the initial state of the FBL. Two constraint equations are given as

$$\begin{aligned} f_3 &\equiv k_1 \cos(\theta_i) + k_2 \cos(\xi_1) - k_3 \cos(\xi_2) \\ &\quad - k_4 \cos(\theta_f) = 0 \\ f_4 &\equiv k_1 \sin(\theta_i) + k_2 \sin(\xi_1) - k_3 \sin(\xi_2) \\ &\quad - k_4 \sin(\theta_f) = 0 \end{aligned} \quad (10)$$

From the equation (10), the equations of two dependent variables θ_f and ξ_1 are derived as

$$\theta_f = \cos^{-1}\left(\frac{k_1^2 - k_2^2 + k_3^2 + k_4^2 - 2k_1 k_3 \cos(\theta_i - \xi_2)}{2k_4 \sqrt{k_1^2 + k_3^2 - 2k_1 k_3 \cos(\theta_i - \xi_2)}}\right) + \chi \quad (11)$$

where $\chi = \tan^{-1} \left(\frac{k_1 \sin(\theta_i) - k_3 \sin(\xi_2)}{k_1 \cos(\theta_i) - k_3 \cos(\xi_2)} \right)$

when $k_1 \cos(\theta_i) \geq k_3 \cos(\xi_2)$, or

$$\chi = \pi + \tan^{-1} \left(\frac{k_1 \sin(\theta_i) - k_3 \sin(\xi_2)}{k_1 \cos(\theta_i) - k_3 \cos(\xi_2)} \right)$$

when $k_3 \cos(\xi_2) > k_1 \cos(\theta_i)$.

$$\xi_1 = \left(\frac{k_3 \cos(\xi_2) + k_4 \cos(\theta_f) - k_1 \cos(\theta_i)}{k_2} \right) \quad (12)$$

In the equation (12), there are two possible values of ξ_1 within domain $0 \leq \xi_1 \leq 2\pi$. Thus, the exact value of ξ_1 has to be selected from satisfying with the equation (10).

3. Sensitivity analysis of the FCM

Variations of dependent variables $\underline{\Omega}: \mathbb{R}^m \rightarrow \mathbb{R}^n$ by design variables $\underline{\Psi}$ can be expressed as

$$\delta \Omega_i = z_{ij}(0) \delta \Psi_j \quad (13)$$

where z_{ij} is a sensitivity coefficient matrix expressing the sensitivity of dependent variables to small changes in design variables.

From ASSUMPTIONS 1 and 2, the analyzing process of z_{ij} is divided into two cases. The one is when $\underline{C3}_u$ and $\underline{C1}_u$ are tangent and the other is when $\underline{C3}_u$ and $\underline{L1}_u$ are tangent

3.1. Sensitivity coefficient matrix when $\underline{C3}_u$ and $\underline{C1}_u$ are tangent

In this case, there are nine design variables $\underline{\zeta} = [l_1, r_1, r_3, D, d_y, k_1, k_2, k_3, k_4]^T$ which

are independent each other, four dependent variables $\underline{w}(\underline{\zeta}) = [\theta_{cl}, \phi_{cl}, \theta_f, \xi_1]^T$ and four

constraint equations $\underline{f}(\underline{\zeta}, \underline{w}) = [f_1, f_2, f_3, f_4]^T$.

Four constraint equations can be differentiated with the design variables $\underline{\zeta}$ by using the chain rule as

$$[u_{mn}] [a_{pq}] + [b_{kl}] = 0 \quad (14)$$

where $u_{mn} = \partial f_m / \partial w_n$ ($1 \leq m \leq 4; 1 \leq n \leq 4$),

$$a_{pq} = \partial w_p / \partial \zeta_q$$
 ($1 \leq p \leq 4; 1 \leq q \leq 9$),

$b_{kl} = \partial f_k / \partial \zeta_l$ ($1 \leq k \leq 4; 1 \leq l \leq 9$), and $[u_{mn}]$ is a nonsingular matrix.

3.2. Sensitivity coefficient matrix when $\underline{C3}_u$ and $\underline{L1}_u$ are tangent

In this case, there are eight design variables $\underline{\rho} = [r_2, r_3, D, d_y, k_1, k_2, k_3, k_4]^T$ which are independent each other, four dependent variables $\underline{e}(\underline{\rho}) = [\theta_{c2}, t_1, \theta_f, \xi_1]^T$ and four

constraint equations $\underline{g}(\underline{\rho}, \underline{e}) = [g_1, g_2, g_3, g_4]^T$.

Here, $g_3 = f_3$ and $g_4 = f_4$.

Four constraint equations can be differentiated with the design variables $\underline{\rho}$ by using the chain rule as

$$[v_{\alpha\beta}] [s_{\gamma\epsilon}] + [h_{\tau\mu}] = 0 \quad (15)$$

where $v_{\alpha\beta} = \partial g_\alpha / \partial e_\beta$ ($1 \leq \alpha \leq 4; 1 \leq \beta \leq 4$),

$$s_{\gamma\epsilon} = \partial e_\gamma / \partial \rho_\epsilon$$
 ($1 \leq \gamma \leq 4; 1 \leq \epsilon \leq 8$),

$h_{\tau\mu} = \partial g_\tau / \partial \rho_\mu$ ($1 \leq \tau \leq 4; 1 \leq \mu \leq 8$), and $[v_{\alpha\beta}]$ is a nonsingular matrix.

4. Results

4.1. Kinematic analysis of the SCM

While F_1 is acting on the feed regulator, θ_f has the linear relation to d as shown in Fig. 6 and the equation (16). In the case of F_2 , the result is similar to the case of F_1 .

$$\theta_f = (-1)^n A d + \theta_{f,0} \quad (16)$$

where $A = 0.137 \text{ rad/mm}$, $\theta_{f,0} = 2.98 \text{ rad}$, which is the rotational angle of feed regulator link when $d = 0.0[\text{mm}]$, $n=1$ when F_1 is acting on, and $n=2$ when F_2 is acting on.

4.2. Sensitivity analysis

In the sensitivity analysis of the dependent variable, the sensitivity coefficient matrix is normalized to show the mutual relation in design variables as

$$s_{ij} = \frac{z_{ij}}{\max(z_{ij})} \quad (17)$$

Because θ_f is the final output of the feed control mechanism (FCM), θ_f is the most important of all dependent variables. Thus, we focus only on the sensitivity of θ_f .

While d is increasing from 0.0 mm to 3.0 mm by 0.5 mm each, components of normalized sensitivity coefficient matrix (NSCM) are shown in Fig. 7 and Fig. 8, and its statistical data are shown in Table 1. Here, $\bar{m}$ is the mean value and σ is the standard deviation.

5. Conclusions

① θ_f has the linear relation to d . Here,

d is the feeding distance of the feed regulator dial shaft and θ_f is the rotational angle of the feed regulator link.

② d have an little effect on the sensitivity of a design value because σ is very small.

③ In NSCM, mean values of the design variables r_1 , r_2 , r_3 and D are two or three times as large as those of the design variables l_1 , d_y , k_2 and k_3 , and are nine or ten times as large as those of the design variables k_1 and k_4 . Therefore the allowable tolerances of the design variables r_1 , r_2 , r_3 , and D have to be applied three times as tight as those of the design variables l_1 , d_y , k_2 and k_3 , and nine times as tight as those of the design variables k_1 and k_4 .

④ In generally, because the design variables of the SCM is two or three times as sensitive as the design variables of the FBL. Therefore, The SCM should be designed more carefully than the FBL.

Acknowledgement

This research was financially supported by Korea Sewing Machine Co., Ltd. and the authors thank President I. C. Park.

References

1. George H. Martin, "Kinematics and Dynamics of Machines", 2nd edition
2. Donald T. Greenwood, "Principles of Dynamics", 2nd edition, Prentice Hall
3. Walter Ledermann, "Handbook of Applicable Mathematics, Volume IV :

Analysis", John Wiley & Sons

4. C. H. Edwards, JR., "Advanced Calculus of Several Variables", Accademic Press

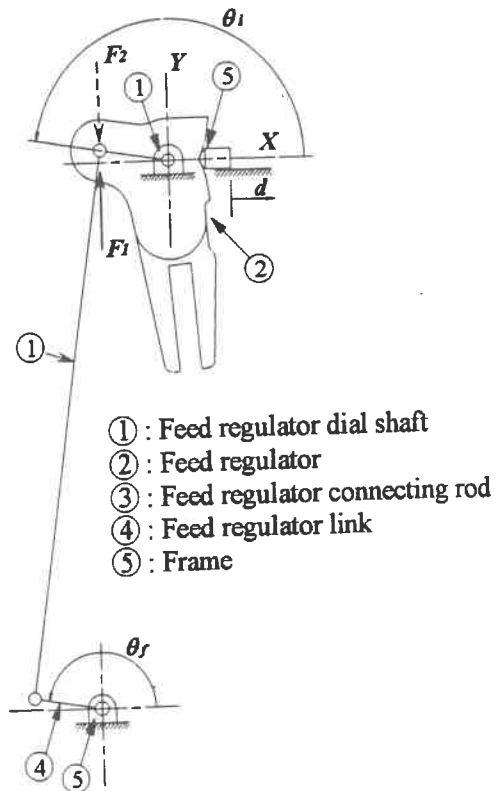


Fig. 1. Feeding Control Mechanism

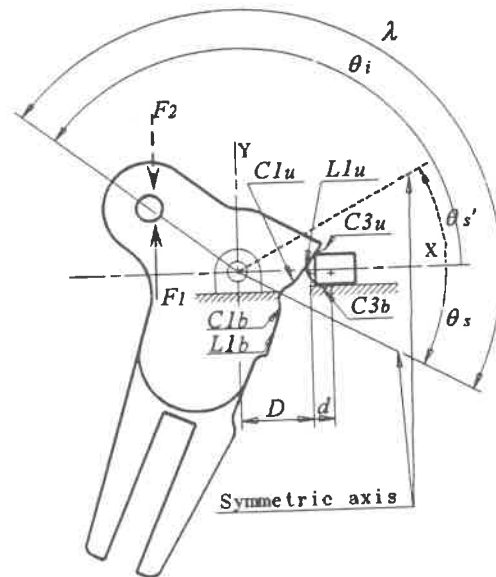


Fig. 2. Sliding contact mechanism when F_1 acts on the feed regulator

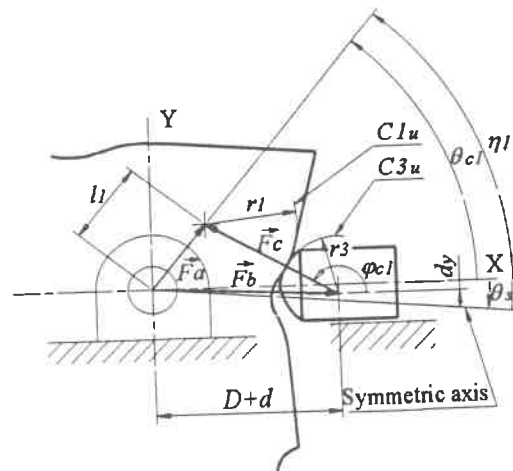


Fig. 3. Kinematic analysis when tangent point is between $C3_u$ and $C1_u$

Acting force	Statistical data	Design variable									
		l_1	r_1	r_2	r_3	D	d_y	k_1	k_2	k_3	k_4
F_1	$\bar{m}$	0.24	1.00	1.00	1.00	-0.87	0.48	-0.09	0.35	-0.35	0.09
	σ	0.098	0.000	0.000	0.000	0.022	0.039	0.060	0.059	0.059	0.060
F_2	$\bar{m}$	-0.29	-1.00	-1.00	-1.00	0.88	-0.48	0.10	0.35	-0.35	-0.10
	σ	0.098	0.000	0.000	0.000	0.022	0.039	0.058	0.057	0.057	0.058

Table 1. statistical data for θ_f

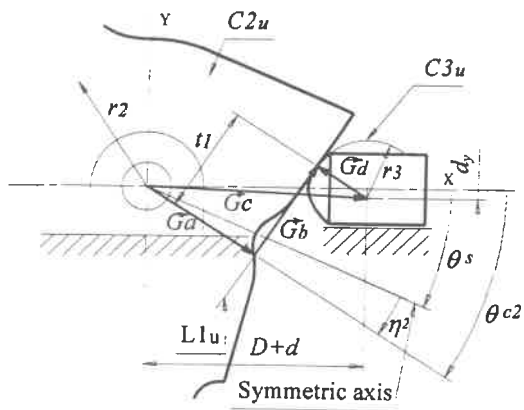


Fig. 4. Kinematic analysis when tangent point is between $C3_u$ and $L1_u$

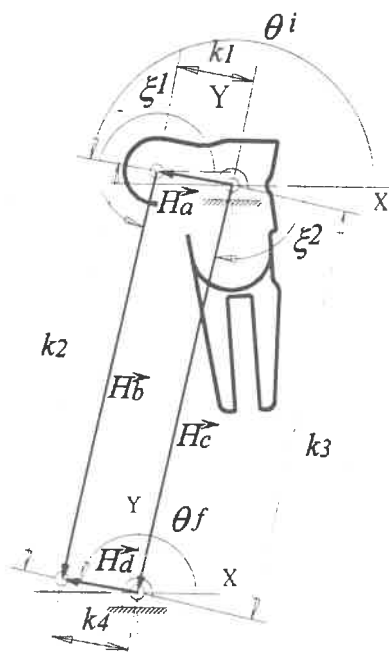


Fig. 5. Four bar linkage

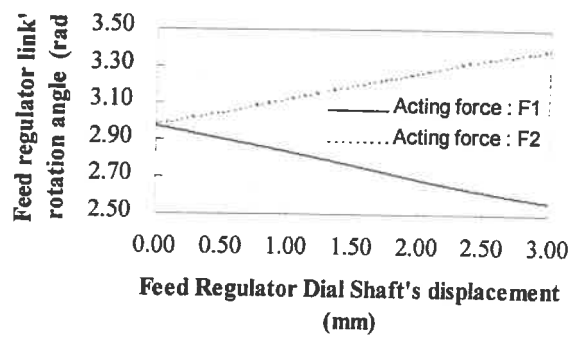


Fig. 6. Relation between d and θ_f

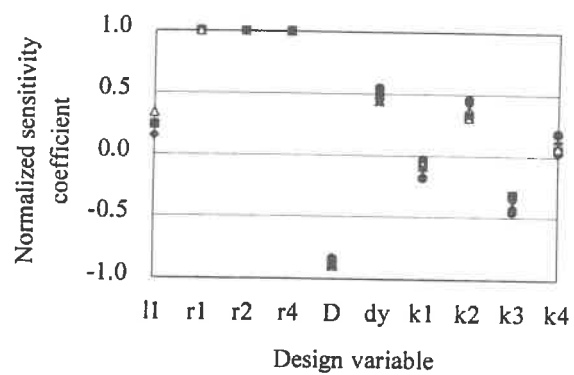


Fig. 7. Component of NSCM when F_1 is acting on

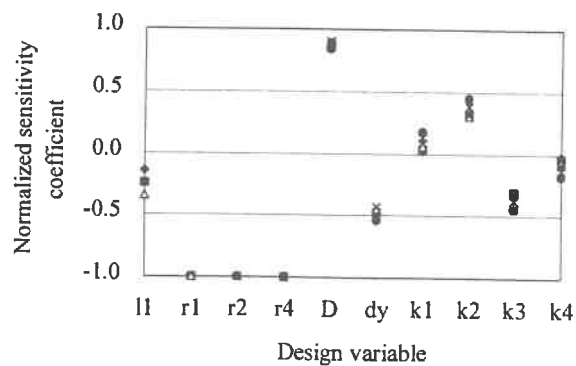


Fig. 8. Component of NSCM when F_2 is acting on

Piezoelectric Microactuator Including Compliant Mechanism

Daniel Enke

Technical University of Ilmenau, Faculty of Mechanical Engineering
Department of Microsystem Technology
PO-Box 0565, 98684 Ilmenau / Germany

Abstract

For the application in optical telecommunications a two dimensional microactuator is developed, which can adjust two optical waveguides to each other. Fibres, integrated optical chips and optical elements like lenses can be positioned. The piezoelectric bending actuator is a planar structure. It can be used for several tasks, e.g. for manipulators, adjusting actuators and microgrippers. The following paper describes the dimensioning, optimization, manufacturing and technical details of the manipulator, with its functional components for movement generation, transmission and its link to the mechanical environment. In the introduction aspects for miniaturizing of bending actuators will be discussed.

Introduction

The core diameter of monomode fibres requires very low mechanical tolerances for coupling. Precise manufacturing essentially determines the costs of connectors (see [1]). The microactuator is one part of an active flexible connector, that means a microsystem, which adjusts two optical waveguides itself. The damping of the coupler will be below 1 dB [2, 3]. Figure 1 shows the 6 axes of misalignment. The x- and y-direction have the biggest influence on the damping of a connector. These are the two axes we adjust. The maximum of movement range ($100\mu\text{m}$), and the accuracy of the actuator ($<100\text{nm}$) follow from the maximum of misalignment and the core diameter. Piezoelectric actuators (like PZT-Ceramics) are often used in mechanical and optical appliances for movements in the micrometer range. You can buy macroscopic appliances for microscopical movements. They are big and heavy.

But piezoelectric actuators are easy to combine with microtechnologies, because of their

- planar structure,
- high energy density,
- very low energy consumption in static running ($<1\text{ mW}$).

The disadvantages are the high driving voltages and the very small deflections with a contraction of less than 0,05% in their length. For micropositioning a bigger motion range and smaller forces are needed. Therefore we use a hinge gear for the movement transmission. Two-axes in motion are realized serially. For the x-direction (out of the wafer) a bimorph is used, the y-direction has elastic hinges in a compliant mechanism.

How to dimension the microactuator

The microactuator has to be investigated with its connection to the surrounding microsystem. The limits of miniaturizing result from:

- feasible energy density,
- feasible deflection and force (stiffness adaption),
- getting losses away from the chip,
- optical limits,
- technological limits.

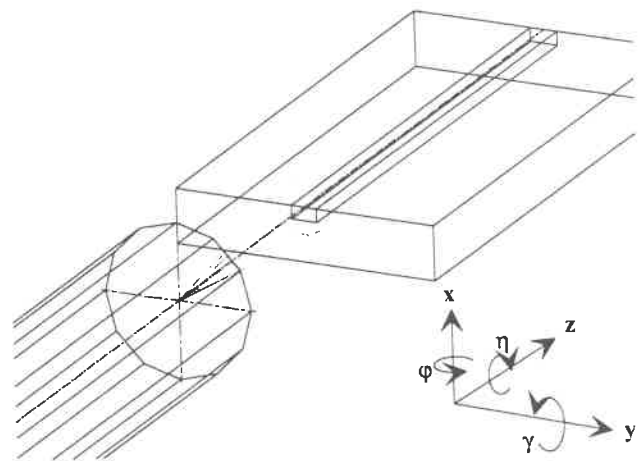


Fig. 1: Adjusting movements in a fibre-chip-coupler

feasible energy density: The volume of a piezoelectric microactuator which is necessary to get out the mechanical energy is calculated from the power balance. As the movement is slow, only the static running is taken into consideration.

$$V_{\min} = \frac{2 \cdot W_{\text{mech}}}{k^2 \epsilon_0 \epsilon_r E_{\max}^2} \quad (1)$$

W_{mech} ... necessary energy
 k ... el.-mech. coupling factor
 $E_{\max}$... max of piezo field strength

The actuating volume sufficient for the maximum bending of the fibre is: $V_{\min} = 0,089 \text{ mm}^3$. That is equivalent to a sheet of piezoelectric ceramics: the size of 1 mm by 1 mm, 100 microns thick.

feasible deflection and force: The forces (F) and displacements (s) are, if the voltage remains constant, in a linear connection with a rise of F_a/s_a , its spring stiffness (fig. 2). In operation with a flexible load $F=f(s)$ (fibre) the operating point OP appears as the intersection of both characteristics. Adaption of stiffness (c) is fulfilled if:

$$c = \frac{F_0}{s_0} = \frac{F_a}{s_a}$$

Thus the maximum deflection and forces are dimensioned to:

$$s_a = 2 \cdot s_0 \text{ und } F_a = 2 \cdot F_0$$

The dimensioning of both gears will be discussed later in this article.

getting losses away from the chip: In the quasistatic running losses are generated only by a leakage current. These losses can be neglected.

optical limits: Because the piezoactuator is polarized, the movement is only a contraction in one direction. The transmitted movement is a rotation, an angular deflection. The maximum angles $\varphi_{\max}, \gamma_{\max} = 1^\circ$ (see fig. 1) may be tolerated by the required damping (see introduction). Thus the minimum radius of the rotation and the minimum bending length are calculated as limits of miniaturizing.

technological limits: There are two kinds of limits caused by manufacturing: fixed and movable. Fixed limits in the processes of silicon micromachining are the minimum ditch widths as a function of the wafer thickness. They

fix the minimum hinge distance to 300 microns. Some limits are movable such as the smallest size of structure (mask and etching technology: 10 to 30 microns). Due to manual assembly we also get a minimum working size.

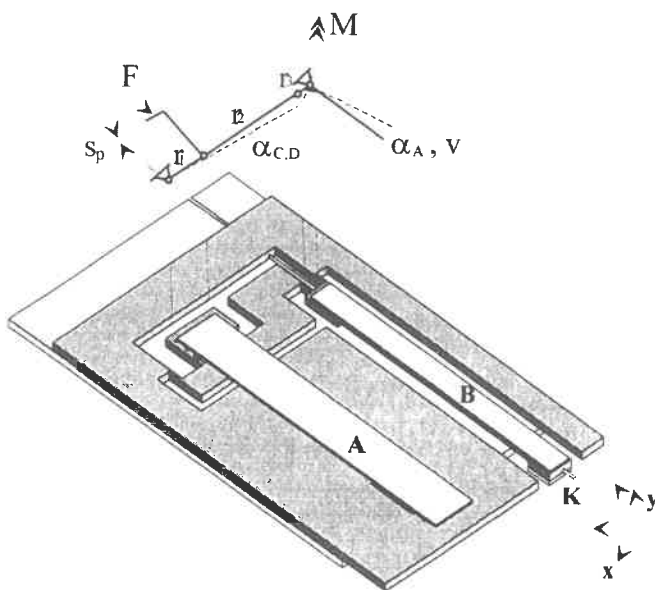


Fig. 3: Schematic display of the piezoelectric microactuator

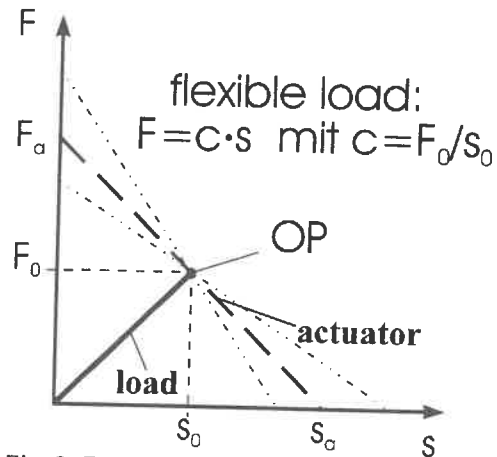


Fig. 2: Force displacement of actuator and load

The compliant mechanism works by the bending of silicon hinges. The piezoceramics plate A contracts (Monomorph) and bows the mechanism in two gear steps, which results in a movement in the plane of the structure (y-direction). The movement straight up from this plane (x-direction) results from a completely glued piezoceramics together with the silicon tongue B (Bimorph). The position of the elastic hinges you can see in the layout in figure 3. In dashed lines you can see the maximum shape of the compliant mechanism. The coupling point K moves into the two adjusting directions.

At first we describe the bimorph system (**x-direction**). The movement from an active layer (piezoceramics plate) is brought into a passive layer (silicon tongue) through a complete connection of the surfaces. The deflection (without load) can be calculated:

$$v(L) = \frac{3}{4} \cdot f_v \cdot d_{31} E \cdot \frac{L^2}{h_{pi}} \quad (2) \text{ from [4]}$$

d_{31} ... piezomodule
 E ... field strength
 h_{pi} ... thickness of piezolayer
 f_v ... constant

The deflection depends on the length L and the thickness of the bimorph tongue. The constant f_v involves the proportions of thickness in the bimorph layers and their Young's moduli. For this actuator $f_v = 0,625$. The thickness of the piezolayer has technological reasons, at this time a minimum of 100 microns. Thus the minimum length $L=10$ mm for the required deflection range. The force in the bimorph tongue depends on:

$$F = \frac{3}{2} \cdot f_m \cdot \frac{d_{31} E}{S_{11}} \cdot \frac{h_{pi}^2 \cdot b}{L} \quad (3) \text{ from [4]}$$

b ... width of the tongue
 S_{11} ... stiffness of the piezoceramics
 f_m ... constant

The usable force of the tongue depends on the momentum deflection. The constant f_m involves the proportions of layer thickness and their Young's moduli. The block force is $F_a = 7,5 \text{ mN}$. From maximum deflection and block force the stiffness of the whole actuator can be calculated, which has to be adapted to the stiffness of the fibre:

$$\frac{F_a}{v_a} = c_{act} = c_{fibre} = \frac{F_{fibre}}{v_{fibre}}$$

The dimensional equation for the adapted bimorph tongue can be written. The left side shows the variable parameters:

$$\frac{f_m}{f_v} \cdot h_{pi}^3 \cdot b = \frac{3\pi}{128} \cdot \frac{S_{pi}}{S_{fibre}} \cdot d_{fibre}^4 \quad (4)$$

For the miniaturizing limits it is important to know that the length of the bimorph tongue (if its equal to the bending length of the fibre) does not take any influence in the stiffness adaption of the x-direction (see equation 4).

The second movement direction in the wafer plane (**y-direction**) uses the bending of microhinges summarized to a flexible mechanism. Two gear steps enlarge the deflection of the piezotranslator and the total transmission rate is:

$$\ddot{u}_{tot} = \ddot{u}_1 \cdot \ddot{u}_2 = \frac{L}{r_1} \left(1 + \frac{r_2}{r_3} \right) \quad (5)$$

The elastic hinges take up energy for their bending (W_M), and the usable energy (W_L) becomes smaller:

$$W_{mech} = W_L + W_M$$

If you take the same transmission ratio in both gear steps the total transmission is on a maximum. But if you take a higher transmission rate in the last gear step, the bending and its energy is smaller. These two arguments have to be adapted to the special task.

The basic equations (deflection, force) express, what the influence of the turning stiffness γ_i and the lengths of levers r_i is like. In equation 6 you can see the case of stiffness adaption, if $c_{act} = c_{fibre}$. In this direction the width of the piezoceramics has an influence on the deflection.

$$\frac{1}{r_1} \left[\gamma_A \cdot \ddot{u}_\alpha + \gamma_B \cdot (\ddot{u}_\alpha + r_1 + 2(r_2 + r_3)) + \frac{\gamma_C + \gamma_D}{\ddot{u}_\alpha} \right] = 3 \cdot \frac{(EI)_{fibre}}{L^2} \text{ with } \ddot{u}_\alpha = \frac{r_2 + r_3}{r_1} \quad (6)$$

The optimization of the hinges in the last gear step (with biggest deflection) is the most important, the turning stiffness of hinges A and B have the strongest influence on the stiffness of the microactuator.

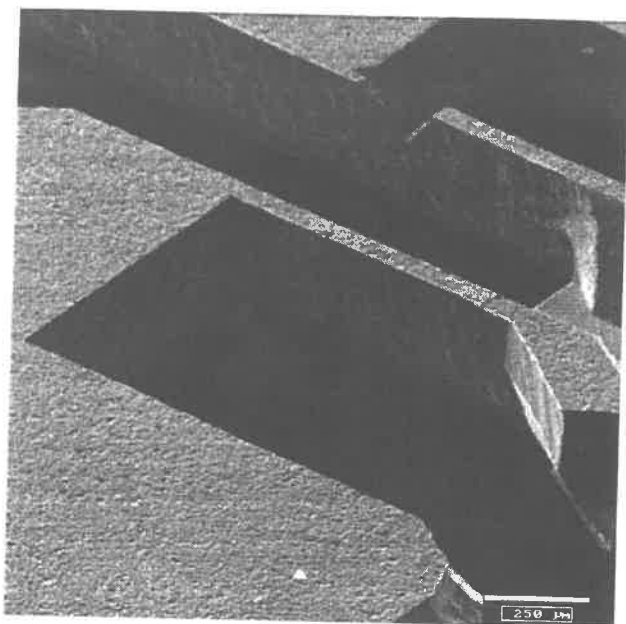


Fig. 5: SEM of a hinge structure

Manufacturing and Measuring

All hinge structures are made in (100)-silicon by wet anisotropic etching on both sides in one etch step. In the cross-section they are nearly rectangular (see fig. 5) with an aspect ratio bigger than 1:5. The fibre will be guided through the actuator between a V-groove on the side below of the chip and a grounding substrate. Thus a z-movement is possible and an adjusting possibility for fibres, keeping the polarization direction, is given in the fibre axis (η). The piezoceramics plates are mounted on the surface of the wafer.

The following table gives technical details:

parameter:	
length, width, height:	15 mm, 10 mm, 1 mm
mass:	100 mg
deflection range:	per axis 100 μ m
input voltage range:	0 - 100 V
positioning time:	2 ms
block force:	10 mN
	no energy consumption in static running; hold position at galvanic separation

Further Development

Further optimization of the micropositioner is directed towards miniaturizing and the use of new technologies in connecting the silicon and the piezoelectric layer. The goals are smaller driving voltages and higher resonance frequencies, because of a smaller mass. Much more uses are possible for this actuator, e.g. microgrippers, fibre-switches and multiplexer, in sensor technology, surface metrology and precision positioning.

Acknowledgement

The work presented is supported by the German Federal Government (contract number: 13MV0316) and the Thuringian State Government (contract number: W6-918/3-183).

References:

- [1] Lutzke, D.: *Lichtwellenleitertechnik : Bauelemente, Systeme, Meßtechnik*. (eng.: Fibretechnics) München: Pflaum, 1986.
- [2] Enke, D.; Frank T.; Gerlach T.: *Mikropositionierantriebe für Lichtwellenleiter* (eng.: Micropositioner for optical waveguides) In: Tagungsband des 41. Int. Wiss. Kolloquium Ilmenau, 23.-26.9.96. Technical University of Ilmenau
- [3] Gerlach, T.; Enke, D.; Frank, T.; Hutschenreuther, L.; Schacht, H.-J.; Schüler, R.: *Towards an integrated microsystem for the automatic adjustment of monomode optical waveguides*. In: Proceedings of the Workshop „MicroMechanics Europe MME'96“. Barcelona 21-22 October 1996
- [4] Pfeifer, G.: *Piezoelektrische lineare Stellantriebe*. (eng.: Piezoelectric linear actuators) Wiss. Schriftenreihe der TH Karl-Marx-Stadt (TU Chemnitz). 6/1982

A SEMI-TOROIDAL FLEXURE SYSTEM FOR EXTENDED LINEAR MOTION

C. Jeffrey Finlayson and James F. Cuttino
The Department of Mechanical Engineering
The University of Alabama

Introduction

A number of linear motion precision devices rely on flexures to allow small displacements with atomic level accuracies. Flat, annular flexural bearings are often used in applications involving axial motion between concentric cylinders. However, a significant disadvantage to these flexures is the limited range of motion imposed by the necessity to keep stresses below the fatigue limit.

A preliminary study of a new flexure geometry has been conducted with the goal of increasing the range of axial displacement of two cylinders while maintaining concentricity. To improve the flexure's limited range, a new cross sectional profile has been developed and analyzed. As shown in Figure 1, the main feature of the flexure is that it resembles a bottom portion of the shell of a toroid.

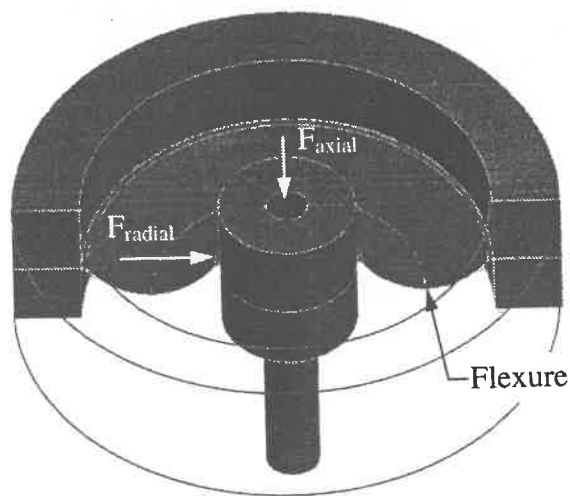


Figure 1: Flexure Configuration

The motivation behind the new geometry was twofold. First, the cross-sectional shape of the flexure allows the flexure to "roll" as the inner cylinder (hereafter referred to as a "shuttle") translates. This rolling action reduces the stresses from bending that are found in the flat flexure. Second, due to its curved shape, the flexure is ideal for conditions where the annular region between the cylinders is pressurized. Since the flexure is curved, it acts as a pressure vessel, reducing bending moments and stresses in the material.

Modeling and Analysis of Flexures

A finite element model of the curved flexure was constructed to analyze advantages and disadvantages of the new design. The study assumed the flexure would be loaded only in the axial and radial directions as shown in Figure 1. Due to symmetry in radial loading, only 1/2 of the flexure was modeled. The flexure's surface was meshed with 6-node, triangular, plate-bending elements. The axial and radial loads were applied separately to examine the two respective stiffnesses. These loads were uniformly distributed over the clamped section in the shuttle. In both loading configurations, the nodes along the outer boundary were constrained in all translations and rotations and the clamped portion in the shuttle was constrained in all rotations. Symmetric boundary conditions were applied along the line of symmetry. Convergence studies were conducted for the models to ensure validity of the data.

Results of Parameter Study

The parameter study was begun by first identifying all physical parameters that characterize the curved flexure bearing. The five principal parameters are the outer radius, r_o , the inner radius, r_i , the radius of curvature, r_c , the fillet radius, r_f , and the thickness, t . These parameters are shown in Figure 2. The distance to the bottom of the flexure is described by a *depth ratio* defined as $(r_o - r_i)/r_c$. This ratio provides a nondimensional parameter that characterizes the depth and curvature of the flexure. This parameter allows the comparison of curved and flat flexures, since $r_c = \infty$ for a flat flexure.

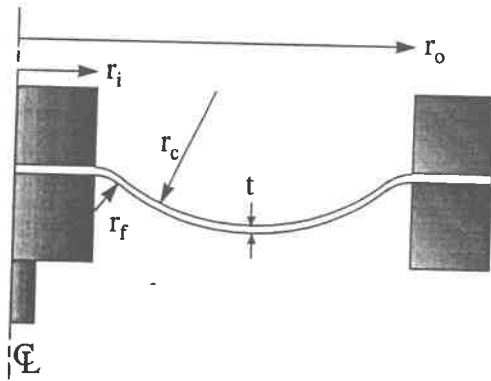


Figure 2: Flexure Parameters

Figure 3 relates stiffness and stress to the principal curvature of the flexure. Figure 3 shows that the highest stress values and lowest stiffness occur at a depth ratio of zero, corresponding to that of the flat flexure. The stress appears to have a “transition” region, dropping off rapidly as the depth ratio is increased to 1 and finally decreasing by a factor of 4 as the depth ratio reaches 2.25. More tests are planned to study this transition region.

The stiffness curve is smoother as the depth ratio is increased. At a ratio of 2.5, the axial stiffness has increased over that in the flat flexure by a factor of 4.5. While the results are not yet conclusive, these preliminary curves show that stress decreases

and axial stiffness increases as the principal curve becomes deeper (decreasing r_c).

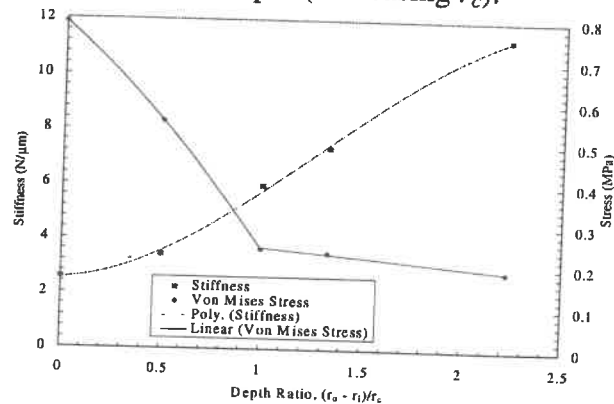


Figure 3: Flexures of Varying Depth Ratios in Axial Loading

Figure 4 plots stress and stiffness versus depth ratio for radial loading. Figure 3 and Figure 4 use the same scales for stress for comparison. In this case, the resulting stress and stiffness curves oppose those in axial loading since the stiffness decreases with increasing depth ratio while stress increases. The results appear to simulate the analogous case of a curved beam under axial loading. Both stiffness and stress show exponential behavior as a function of depth ratio. The stress value for radial loading is only 3% of that for axial loading of the flat flexure, but it increases to 90% for the maximum depth ratio. This demonstrates that stresses from radial loads can be as significant as stresses from axial loads for the larger values of depth ratio or can be negligible for small depth ratios.

Figure 5 presents a plot of stiffness and stress versus the ratio of r_o/r_i for flexures in axial loading. Stress and stiffness decrease with increasing r_o/r_i ratios. Stress decreases by 17% as r_o/r_i is increased from 3 to 6. However, stiffness decreases by a factor of almost 2.5 over the same range of r_o/r_i .

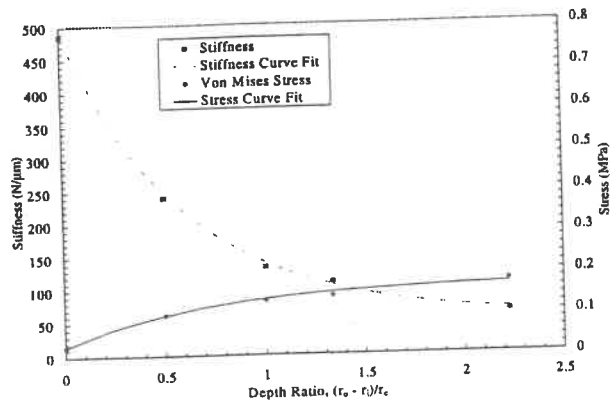


Figure 4: Flexures of Varying Depth Ratios in Radial Loading

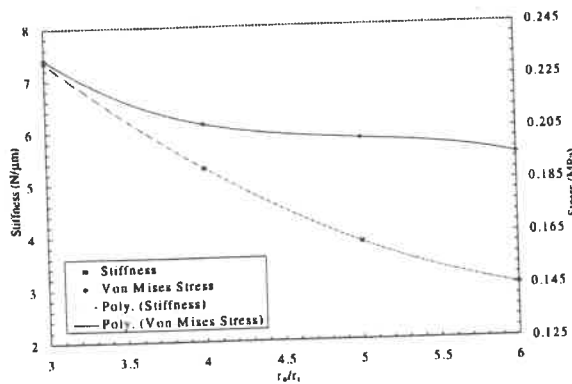


Figure 5: Flexures of Varying r_o/r_i Ratios in Axial Loading

Figure 6 shows stiffness and stress plotted against r_o/r_i ratios for radial loading. Radial stiffness decreases and stress increases with larger r_o/r_i ratios. Stiffness decreases by a factor of 2.7 for an increase of r_o/r_i from 3 to 6. Radial stiffness follows a power relation proportional to $(r_o/r_i)^{1.5}$. Stress increases by a factor of over 1.5 over the same range of r_o/r_i . Figure 5 and Figure 6 have the same stress scale to allow easy comparison. For a r_o/r_i ratio of 3, stresses from axial loading are over 1.6 times larger than stresses from radial loading. However, for a r_o/r_i ratio of 6, stresses from radial loading are slightly larger than the stresses from axial loading.

Figure 7 shows stress and stiffness plotted against flexure thickness under axial loading. Stiffness increases and stress decreases as

thickness is increased. Stiffness appears to be proportional to thickness squared.

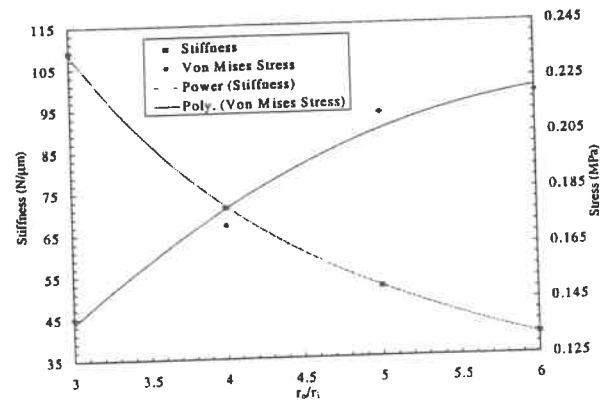


Figure 6: Flexures of Varying r_o/r_i Ratios in Radial Loading

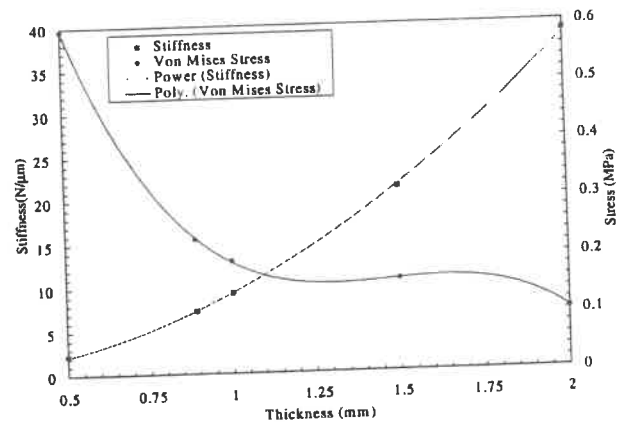


Figure 7: Flexures of Varying Thickness in Axial Loading

In Figure 8 stiffness and stress are plotted against flexure thickness for radial loading. Stiffness is shown to increase and stress decrease with increasing thickness. Stress and stiffness are both a function of thickness raised to a power similar to the axial loading case.

Figure 9 displays stiffness and stress plotted against fillet radius for axial loading. The graph shows maximum stress values decreasing linearly with increasing fillet radii. Stress decreases by 18% over the range of fillet radius of 1 mm to 4 mm. Since the purpose of the fillet is to reduce stresses, this result is expected.

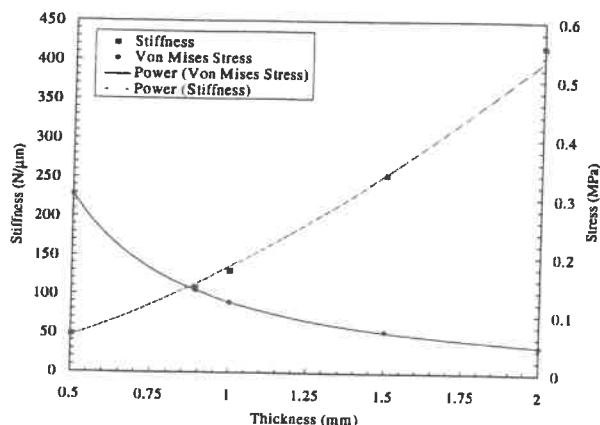


Figure 8: Curved Flexures of Varying Thickness in Radial Loading

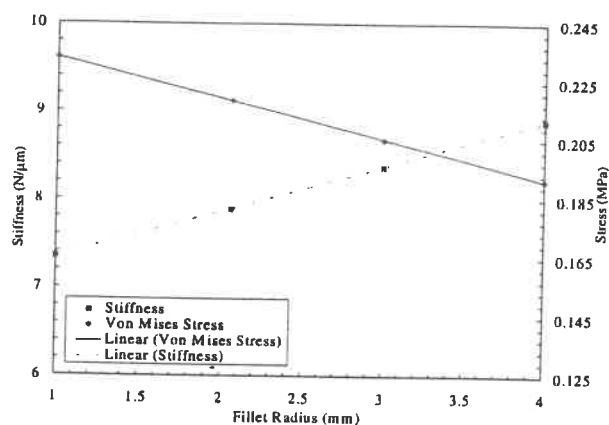


Figure 9: Curved Flexures of Varying Fillet Radii in Axial Loading

The axial stiffness increases linearly by 21% as fillet radius is increased from 1 mm to 4 mm. Whereas the axial stiffness increased for a decreasing radius for the principal curve, it increases for *increasing* fillet radius. Although this effect may be due to the shortening of the principal curve, further study is planned to isolate the cause. It can be concluded, however, that median values of fillet radii offer a compromise between low axial stiffness and low stress levels. For actuators with high force to displacement ratios, the larger fillet radii are optimal.

Figure 10 shows that the stress values for radial loading are less pronounced than for axial loading, changing by 9%. As expected, the radial stiffness increases as the fillet radius

is increased. The stiffness also varies by 9% over the range of fillet radii.

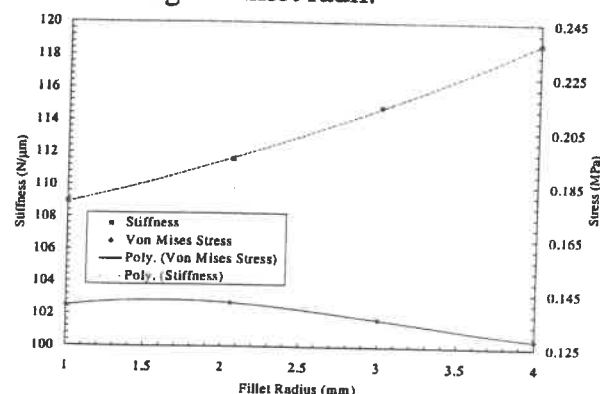


Figure 10: Curved Flexures of Varying Fillet Radii in Radial Loading

Conclusions

Based on the new geometry, a curved flexure stressed to the same stress level as a flat flexure provides over 4 times more deflection. However, radial stiffness decreases by a factor of 5. For cases where the radial direction is the non-sensitive direction, the gain in axial displacement may justify the reduction in radial stiffness.

The results show that for axial loading, stiffness increases and stress decreases with increasing depth ratios, flexure thickness, and fillet radius. Stress and stiffness decrease with larger r_o/r_i ratios. For radial loading, stiffness increases with increasing fillet radii and thickness, but decreases with increasing depth ratios and r_o/r_i ratios. Stress decreases with increasing fillet radius and thickness, but increases with larger depth ratios and r_o/r_i ratios.

Preliminary work has begun on a dimensional analysis to identify functional relationships between stiffness, stress and the geometric parameters presented in this study. Based on this work, a design methodology for toroidal flexures will be developed.

MICROSCOPIC BEHAVIOR OF LEADSCREW-NUT SYSTEM AND ITS IMPROVEMENT

Shigeo Fukada
Department of Mechanical Systems Engineering
Shinshu University
500 Wakasato, Nagano 380 JAPAN

1. Introduction

The positioning accuracy of a precision stage driven by a leadscrew-nut system depends mostly on the characteristics of the leadscrew in microscopic range. Recently it has been suggested that preloaded ball screws have some nonlinear elastic characteristics between the small torque input and the nut displacement within microscopic range, and that these characteristics are readily available in the nanometer positioning control^{1) 2) 3)}. This paper presents the experimental results on this microscopic behavior of a ball screw and a conventional sliding leadscrew. It also proposes a method of improving the positioning performance of the sliding leadscrew.

2. Microscopic behavior of ball screw

2.1 Experimental method

The experimental apparatus is shown in Fig. 1. The outer diameter of the screw shaft is 32mm and the lead is 5mm/rev. The circuit of the ball circulation is 2.5×2 . The ball screw is double nut type, and the preload is set between the nuts by a specially designed device which consists of adjusting screws and load cells inserted between the nuts. The screw shaft is fixed horizontally, and a small circumferential torque is applied to the nut by solenoids.

The axial displacement and the rotational angle of the nut are measured by capacitive gap sensors. The nut in the setup does not touch the sensors or actuators, and it keeps contact only with the screw shaft through the circulating balls.

2.2 Results and consideration

Fig. 2 shows the quasi-static response of the displacement and the rotational angle of the nut to the applied torque in clockwise (cw) direction when the preload is 900 N. First, the displacement and the rotational angle increase according to the increase of the torque, and next they decrease following the decrease of the torque. This means that the response shows nothing but the elastic characteristics within microscopic range. Fig. 3 shows the relation

between the displacement and the rotational angle. The broken line in the figure represents the theoretical line estimated from the lead of the screw. In this microscopic range, the relation between the displacement and the rotational angle is also subject to the lead of the screw as it is in the macroscopic range.

Fig. 4 shows the relation between the applied torque and the displacement, i.e. stiffness curve, arranged from data similar to Fig. 2 in the case of 600 N preload. The curve shows nonlinear elastic characteristics with hysteretic effect in the cw direc-

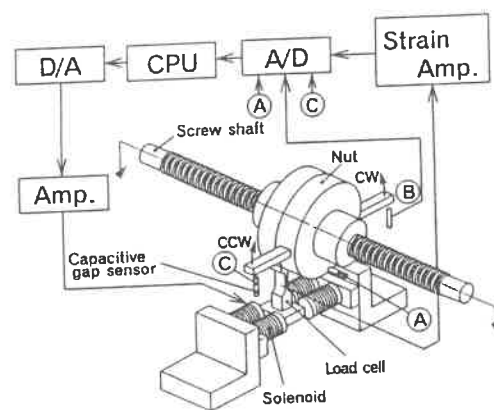


Fig. 1 Experimental apparatus for ball screw

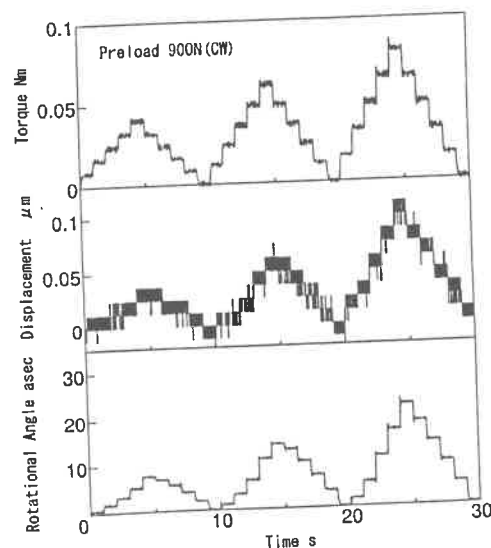


Fig. 2 Response of displacement and rotational angle

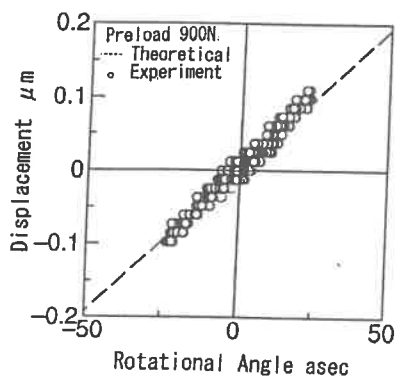


Fig. 3 Displacement vs. rotational angle

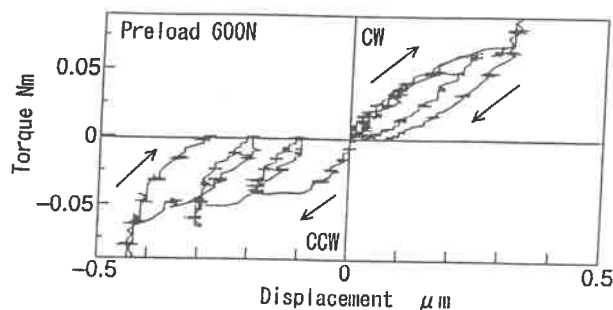


Fig. 4 Torque vs. displacement

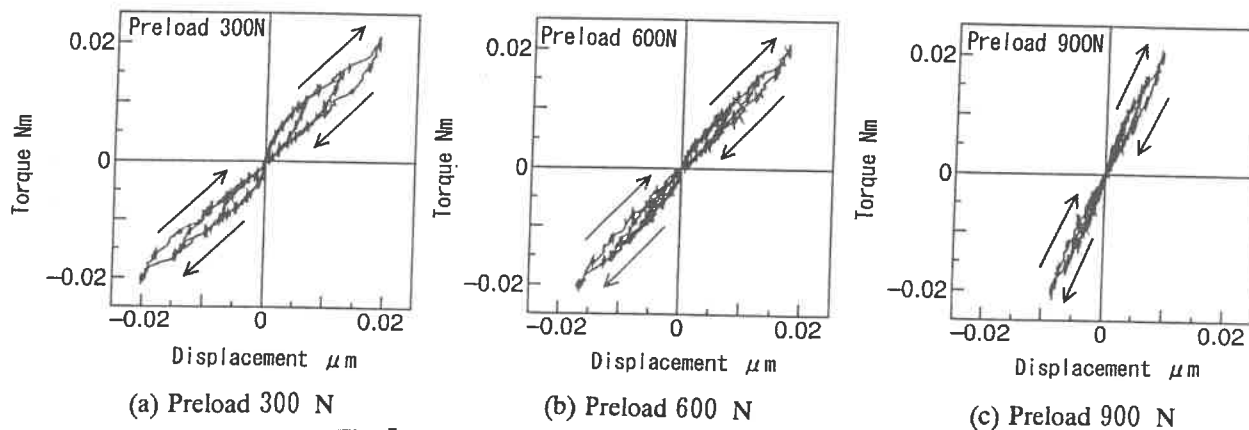


Fig. 5 Stiffness curves

tion. In the counter-clockwise (ccw) direction, the torque exceeds the limit of elastic range at -0.05 Nm, and the nut rotates macroscopically.

Further experiments are performed in more microscopic range with higher resolution. Fig. 5 shows the stiffness curves when the preload is changed into three values. The elastic characteristics becomes stiffer and more linear according to the increase of the preload.

3. Characteristics of positioning table driven by sliding leadscrew

3.1 Experimental method

To consider the characteristics of sliding leadscrew, some experiments are performed on a positioning table driven by a sliding leadscrew as shown in Fig. 6. The table is guided by a cross-roller guide, and the stroke of the table is 100 mm. The sliding leadscrew has a trapezoidal screw thread of $\text{Tr}14 \times 2$. A small preload is given between double nuts to eliminate backlash. The leadscrew is coupled to a DC servo motor of 80 W through coupling. The motor is driven by a linear current amplifier. The

table position is detected by a optical linear encoder of 10 nm resolution, and can be controlled by software servo system with full closed feedback loop.

3.2 Results and consideration

As with the experiments on the ball screw, the response of the table displacement to the torque input is investigated. Fig. 7 shows changes of the torque and the table displacement and their relationship. Unlike the case of ball screw, the table displacement stays in a state of complete rest until the motor torque exceeds a certain amplitude. The preload of the leadscrew was adjusted before assembling so that the start torque is 0.1 Nm, and the figure shows the table starts with the start torque of the leadscrew.

Fig. 7 has the same dimensions as Fig. 4. The figure shows no elastic characteristics, and the table is stationary until it moves macroscopically. This means that the microscopic behaviour of the sliding leadscrew is ruled by Coulomb's friction: this friction spoils the positioning performance of the sliding leadscrew. The positioning accuracy of the table controlled by simple proportional controller is no more than $30 \mu\text{m}$ dispersion.

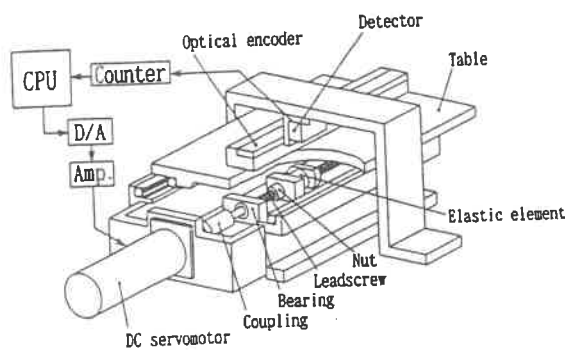


Fig.6 Experimental setup of positioning system using sliding leadscrew

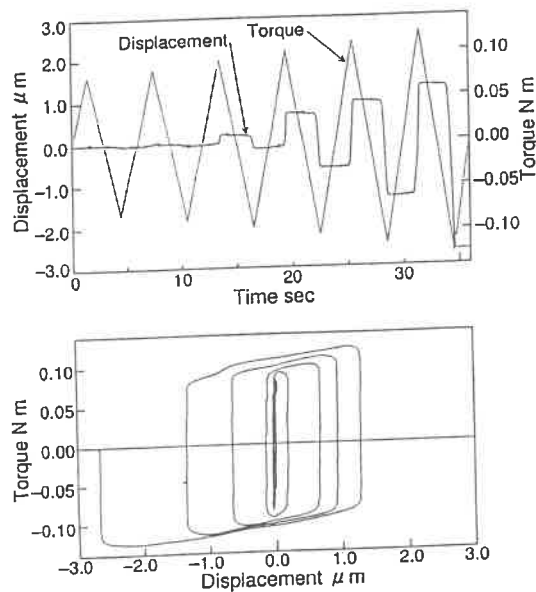


Fig.7 Torque vs. displacement

4. Improvement of positioning performance of sliding leadscrew

4.1 Principle and design of mechanism

Ball screws have elastic characteristics in the relation between torque input and nut displacement though hysteretic. It has been indicated that this property is very available to the positioning control because the displacement shows no stick and responds to the torque input however small¹⁾. If such elastic characteristics can be given to a sliding leadscrew, its positioning performance will be improved. Considering this, a simple method is proposed.

Fig.8 shows the principle of the proposed method. When the nut is fixed to the screw shaft so that the friction does not exceed the maximum static friction, any torque input to the screw shaft is transmitted to the nut direct. Therefore, if the elastic element which converts the torsional moment into the axial

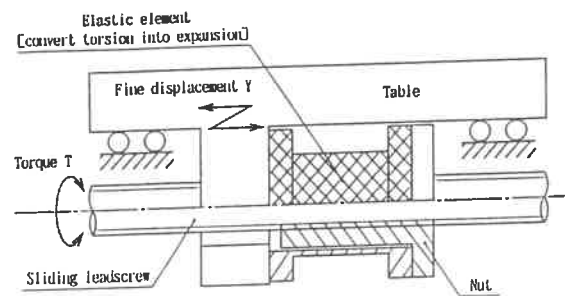


Fig.8 Principle of proposed method

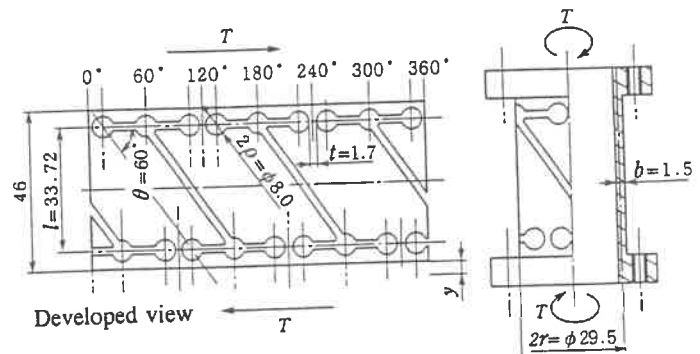


Fig.9 Configuration of elastic element

expansion elastically is inserted between the nut and the table, the table displacement is easily controlled in microscopic range by operating the torque input.

For the elastic element, a cylinder with some flexure hinges formed by notching, as shown in Fig. 9, was devised. When the torsional torque is applied to both ends, the hinges deform elastically, and both end surfaces expand relatively. The relation between the torque T and the expansion y is summarized in the following equation⁴⁾.

$$T = K \cdot y = \frac{4 E b t^{5/2} r \tan \theta}{3 \pi \rho^{1/2} l^2} y \quad (1)$$

where E is Young's modulus and b, ρ, t, l, θ, r are geometrical parameters as shown in Fig. 9. On the basis of the experimental results in the preceding chapter, the stiffness K in eq.(1) can be calculated as $0.1 \text{ Nm}/30 \mu\text{m}$. The dimensions are designed as shown in Fig. 9.

4.2 Experiments and evaluation

The designed elastic element is manufactured using phosphor bronze, and attached to the leadscrew in the system of Fig. 6. Fig. 10 shows the relation between torque input and table displacement. It is confirmed that the torque is converted into the displacement elastically.

The PTP positioning for the step length of 10 mm is performed with PI controller. The control action is

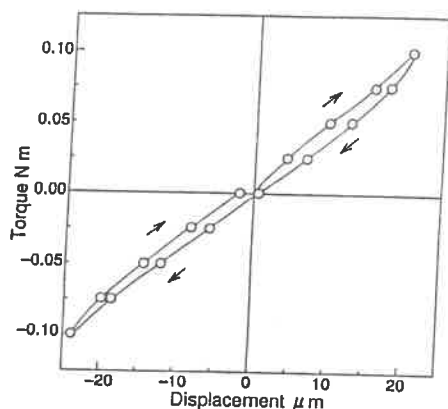
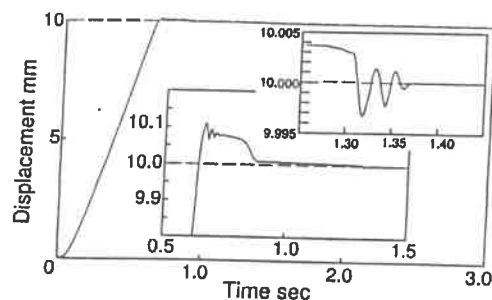


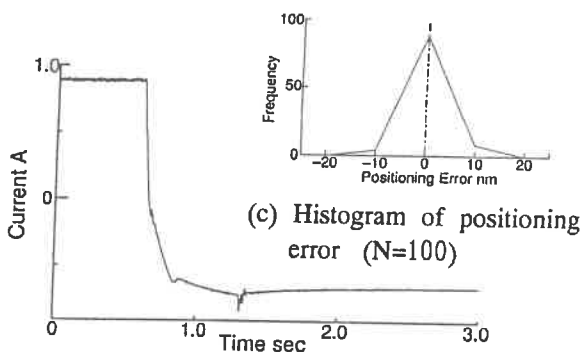
Fig. 10 Torque vs. displacement

switched for the ranges of deviation divided into three regions — coarse, fine and ultra-fine; and the PI parameters are determined for each region by trial-error method. Fig. 11 shows the transition of displacement and motor current during the positioning. The control parameters are switched from coarse to fine at 0.3 mm deviation, and from fine to ultra-fine at 3 μm deviation. The motor current corresponds to the displacement simply and converges smoothly. Fig. 11(c) shows the histogram of the positioning error attempted 100 times. Each error is less than 10 nm which is the resolution of the linear encoder.

Fig. 12 shows the response of the positioning system to 10 nm step input. Fig. 12(a) shows the feedback signal from the linear encoder, and (b) shows the measured data by another capacitive gap sensor independent of the control system. It is confirmed that the positioning system actually responds to an input of 10 nm resolution.

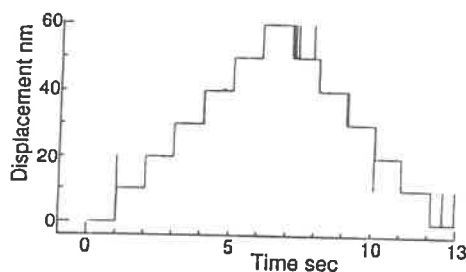


(a) Displacement

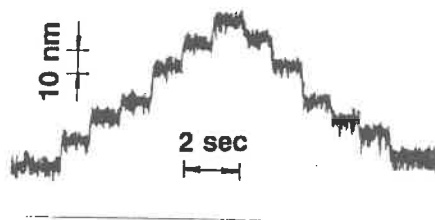


(c) Histogram of positioning error (N=100)

(b) Motor current
Fig. 11 Response to 10 mm step input



(a) Output of linear encoder



(b) Measurement by capacitive gap sensor
Fig. 12 Response to 10 nm step input

5. Conclusions

The conclusions are summarized as follows.

- (1) The experiments on the ball screw show the elastic characteristics in the relation between torque input and nut displacement within microscopic range. The linearity of the elastic property improves when the preload is increased.
- (2) It is shown by another experiment that the sliding leadscrew does not have any elastic characteristics at all.
- (3) A simple mechanism using an elastic element to convert torsional moment into expansion is proposed. It is then verified experimentally that the device is readily available to vastly improve the positioning performance of the sliding leadscrew.

References

- 1) Jiro Otsuka, Shigeo Fukada et al.: Ultraprecision positioning Using Lead Screw Drive, *Int. J. JSPE*, 27, 2 (1993) 142.
- 2) Jiro Otsuka: Nanometer Positioning Using Three Kinds of Leadscrews, *Nanotechnology*, 2, 3 (1992) 29.
- 3) James F. Cuttino and Thomas A. Dow: The Preloaded Ball Screw As A Nanometer Actuator,
- 4) J. M. Paros: How to Design Flexure Hinges, *Machine Design*, Nov. 25 (1965) 151.

Recent Advancements In Piezoelectric Stepping Motors

David Henderson, Dana Jensen and Peter Piccirilli
Burleigh Instruments, Burleigh Park, Fishers, NY 14453

Piezoelectric (PZT) stepping motors such as the Inchworm® combine the high resolution and stiffness of piezoelectric stacks with the long travel of more conventional electromagnetic motors and leadscrews. Recent advances in co-fired PZT materials make it possible to produce a high performance Inchworm that operates at only 60 volts (1). We call this new "inside-out" and spring loaded design the Inchworm II™. This linear motor concept has been further improved to operate at "constant velocity" (2) with minimal "glitch" and achieve continuous rotary motion.

Constant velocity and minimum glitch performance are achieved using a modified Inchworm II design. Previous PZT stepping motors produce discrete motion steps with a finite dwell period while the clamps are changing as shown in Figure 1. The dwell period occurs during each clamp change when both clamps are activated (i.e. both clamps are supporting the moving structure but are stationary). This period of stationary clamp overlap is needed to prevent slippage and loss of actuation force. A constant velocity motor independently moves each clamp-extension assembly and allows the clamps to be moving during the clamp change. This effectively eliminates the stationary dwell period as shown in Figure 2.

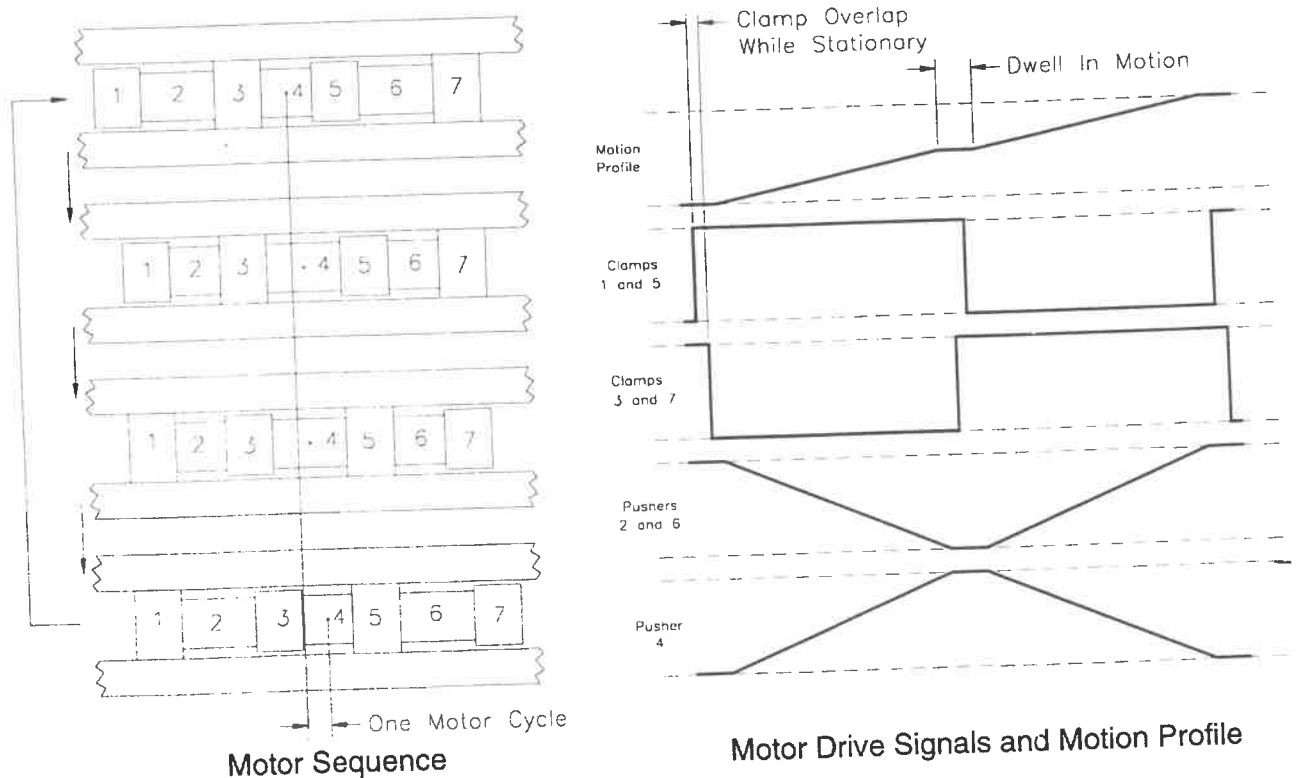


Figure 1 - Clamps Overlap While Stationary And Produce A Dwell In The Motion Profile In Current Inchworm Motors

"Glitch" during clamp changes is minimized by using only the D_{33} motion of the PZT stacks and not the D_{31} motion. The clamping mechanism must produce minimum axial motion

(i.e. motion parallel to the motor's output axis) when a clamp is activated or deactivated. Designs that ignore cross coupling effects produce glitches that are one third of the clamping stroke. (For example a three micrometer clamping stroke produces a glitch of one micrometer.) The Inchworm II glitch is reduced to less than 30 nm using line contacts on the clamp pads and a symmetric configuration as shown in Figure 3.

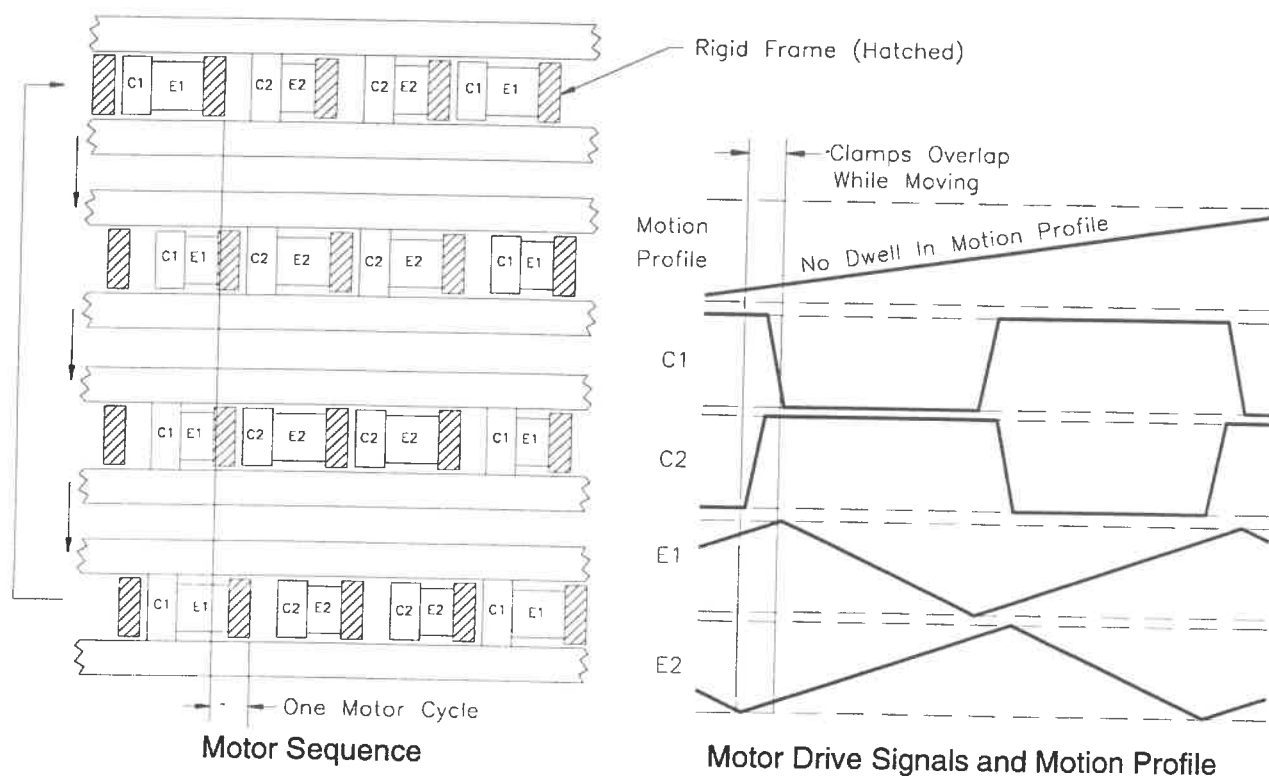


Figure 2 - Clamps Overlap While Moving And Produce No Dwell Using Constant Velocity Approach

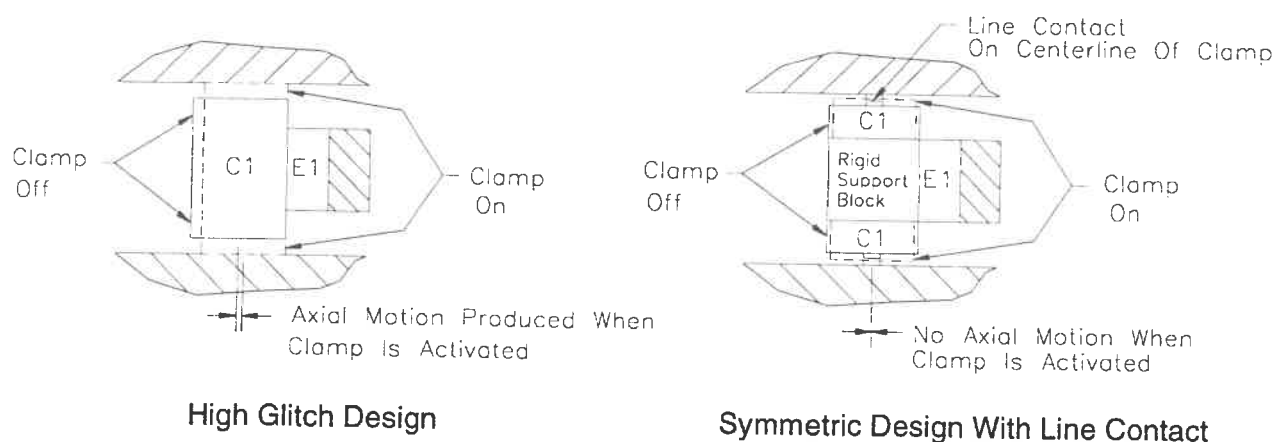


Figure 3 - Proper Clamp Design To Minimize Glitch

The constant velocity and low glitch innovations have been demonstrated with a small stage shown in Figure 4. This stage uses a moving carriage to preload the rails against the

motor assembly. A separate set of rolling element bearings guide the top and bottom plates of the stage.

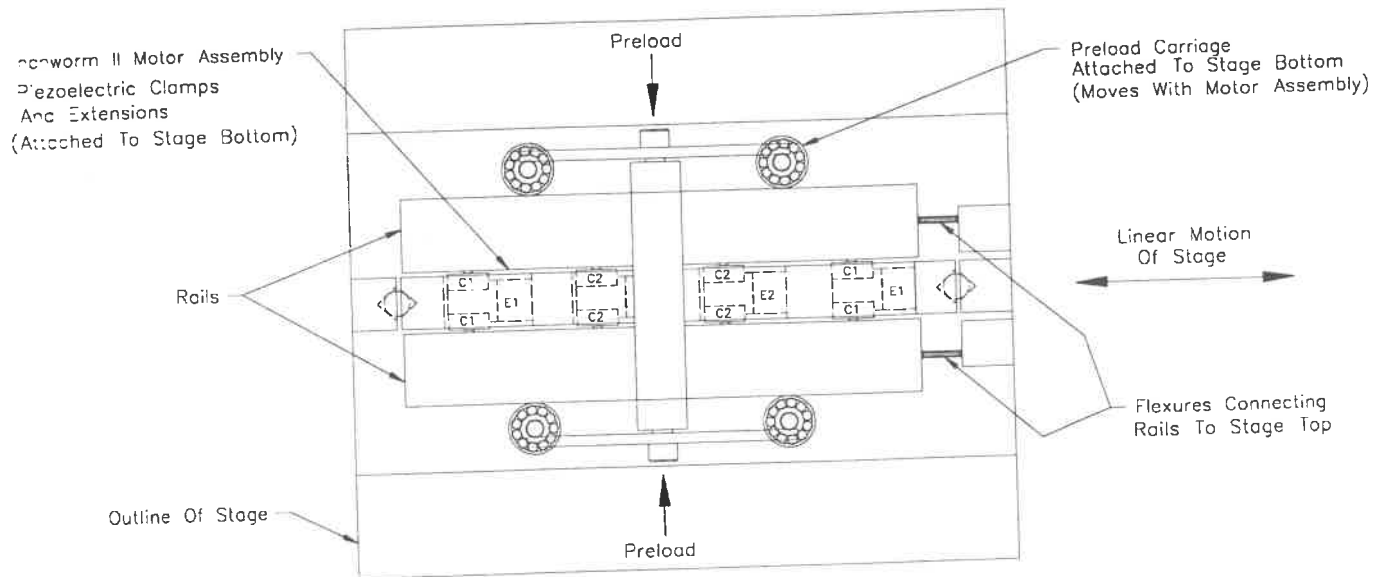


Figure 4 - Prototype Constant Velocity Inchworm II Linear Stage

Demonstrated Prototype Performance

Preload	25 pounds
Output Force	3 pounds
Static Holding Force	10 pounds
Maximum Speed	1 mm/second
Glitch During Clamp Changes	Less than 30 nm (At the resolution limit of test equipment.)

Future Work

- Evaluate the motion profile of this prototype with a higher resolution position encoder to measure the glitch more precisely.
- Simplify the prototype design so that fewer parts are required and manufacturing costs are reduced.
- Optimize the stiffness of the motor assembly frame to reduce stored energy and the corresponding glitch. As the clamps change the load paths in the frame change the energy stored in the frame is released to produce glitch. At the nanometer level this is a significant effect that must be minimized.

A rotary stepping motor has also been designed that uses the constant velocity concept as a tangential drive. Lapped rails are replaced with lapped rings. Six clamp-extension assemblies are used (instead of four) and are sandwiched between two drive rings with designations C1, C2, E1, and E2.. The drive rings rotate and the clamp-extension assemblies are fixed to the housing. The drive rings are pressed against the clamp pads using symmetric opposing forces that are internally supported by the shaft. See Figure 2 for the operating sequence. This new rotary motor design is simple and inexpensive to manufacture. This

concept is also small, stiff, lightweight easily scaled to accommodate larger diameters and hollow shafts.

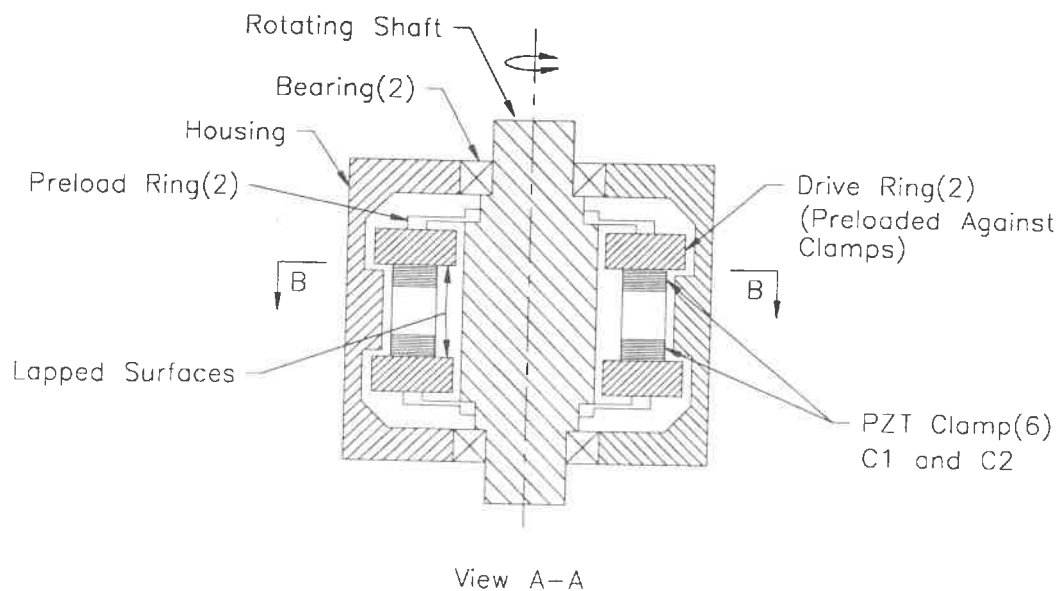
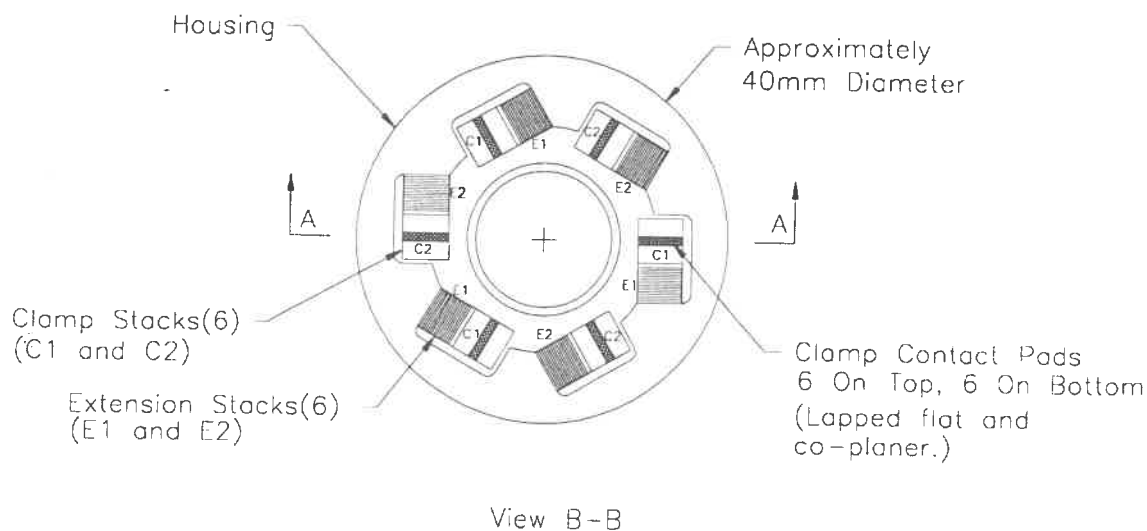


Figure 5 - Cross Section Of Rotary Motor



**Figure 6
Plan View Of Rotary Motor**

References:

1. David Henderson, "A New Inchworm® Motor For Industrial Nanopositioning Applications", ASPE Annual Meeting, 1995.
2. Dan E. Luttrell, Thomas A. Dow, "Design of a Constant Velocity Micro-Stepping Actuator", Precision Engineering Center Annual Report, 1987.

A Long-Range Scanning Stage Design (The LORS Project)

Mike Holmes, Robert Hocken

Precision Engineering Group, University of North Carolina,
Charlotte, NC, USA

David Trumper

Mechanical Engineering Department, Massachusetts Institute of Technology,
Cambridge, MA, USA

This extended abstract describes a magnetically-suspended six-degree-of-freedom precision motion control stage which is presently being built in the Precision Engineering Laboratory at the University of North Carolina at Charlotte. This Long-Range Scanning (LORS) stage will have travel in the horizontal plane of 25 mm² along with 100 μ m of vertical travel. Vertical position feedback will be provided by three capacitance probe sensors while heterodyne laser interferometry will be used for lateral position feedback. The stage is to have a positioning resolution of 0.1 nm, positioning repeatability of 1 nm, and a positioning accuracy of 10 nm. These performance objectives have been chosen to match the measurement requirements associated with present and future production needs for devices such as integrated circuits, photo-masks, and micromechanical actuators.

Keywords: magnetically suspended; precision motion control

Introduction

The objective of this research is to analyze, design, construct, and test a precision motion control stage with performance objectives of 0.1 nm positioning resolution, 1 nm repeatability, and 10 nm accuracy inside the 25 x 25 x 0.1 mm³ range of travel. The LORS stage will be used for the positioning of samples in scanned probe microscopy. Scanned probe microscopy will ultimately be used to characterize the performance of the LORS stage. The LORS stage is scheduled to come on-line in February of 1997 with a project completion date of August 1997. This extended abstract presents a summary of the final design along with some performance predictions.

Mechanical Design

The LORS stage, shown in Figures 1 and 2, consists of a moving element (**platen**), a **machine frame**, and a **metrology head**. The bottom of the platen holds four permanent magnet arrays. A **reference block** is mounted to the top of the platen via three flexures^{1,2}. As in previous work^{3,4}, the platen is light-weighted and floated in oil to reduce bias currents in the motors and to provide viscous damping and high-frequency coupling. The stator windings are mounted in the machine frame opposite the magnet arrays. The reference block contains the reference surfaces for the position feedback sensors, namely, target mirrors for the interferometers and a conductor plane which acts as the targets for the capacitance probes. The sample to be imaged is supported by a sample holder

which is kinematically mounted to the reference block. The metrology head is kinematically mounted to the top of the machine frame. The metrology head contains the feedback sensors, three interferometers for lateral position feedback and three capacitance probes for vertical position feedback. The metrology head also provides support for the scanned probe microscope head. To reduce the effects of thermal drift, the metrology head and the reference block are made of Zerodur, with a coefficient of thermal expansion of⁵ $0.05 \mu\text{m}/\text{m}/^\circ\text{C}$. The idea is to control the currents in the motors, thereby controlling the forces exerted on the platen, based on the position feedback from the interferometers and the capacitance probes. We will be able to image various samples which will be placed on the sample holder. The scanned probe interface mount is designed to support various types of scanned probe microscope heads (i.e., Atomic Force Microscopes (AFM) and Scanning Tunneling Microscopes (STM)).

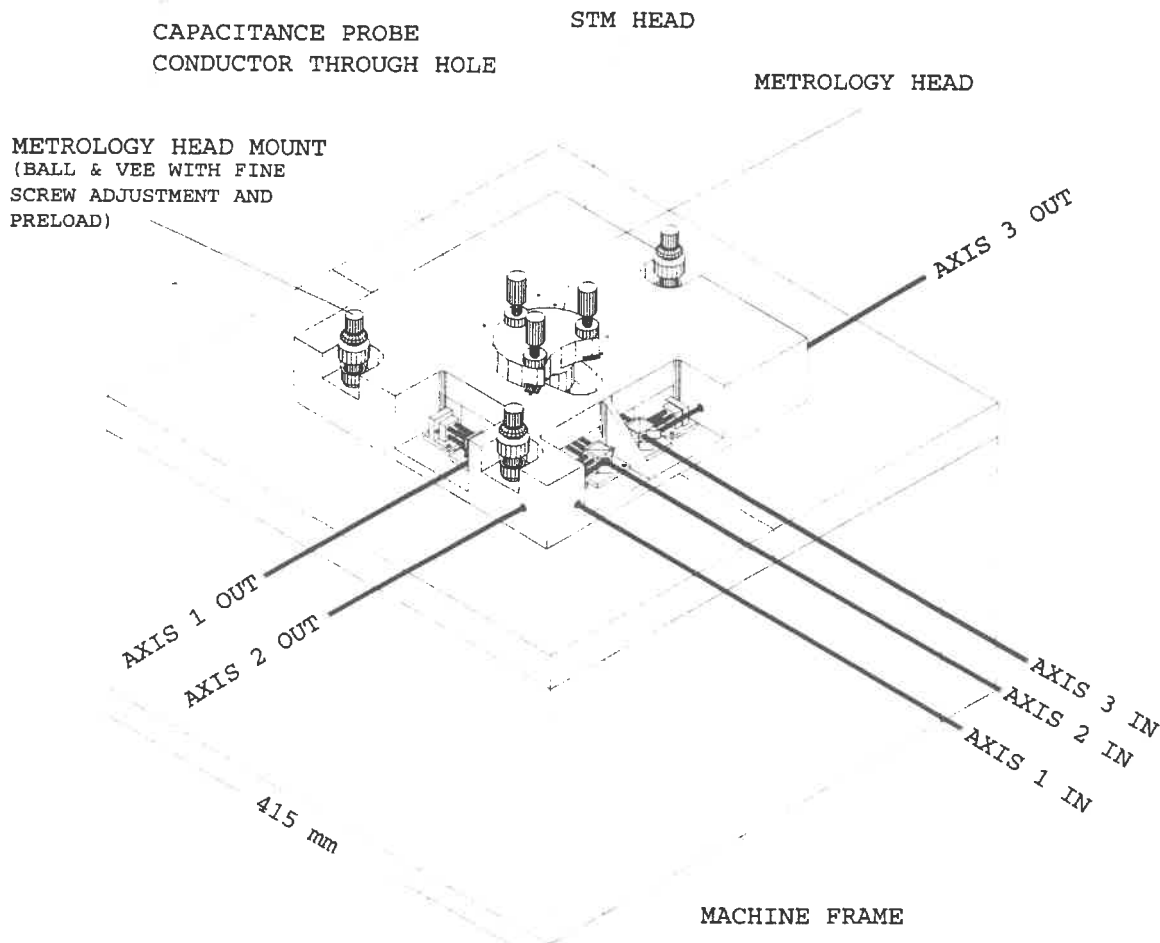


Figure 1: The LORS Stage (External View)

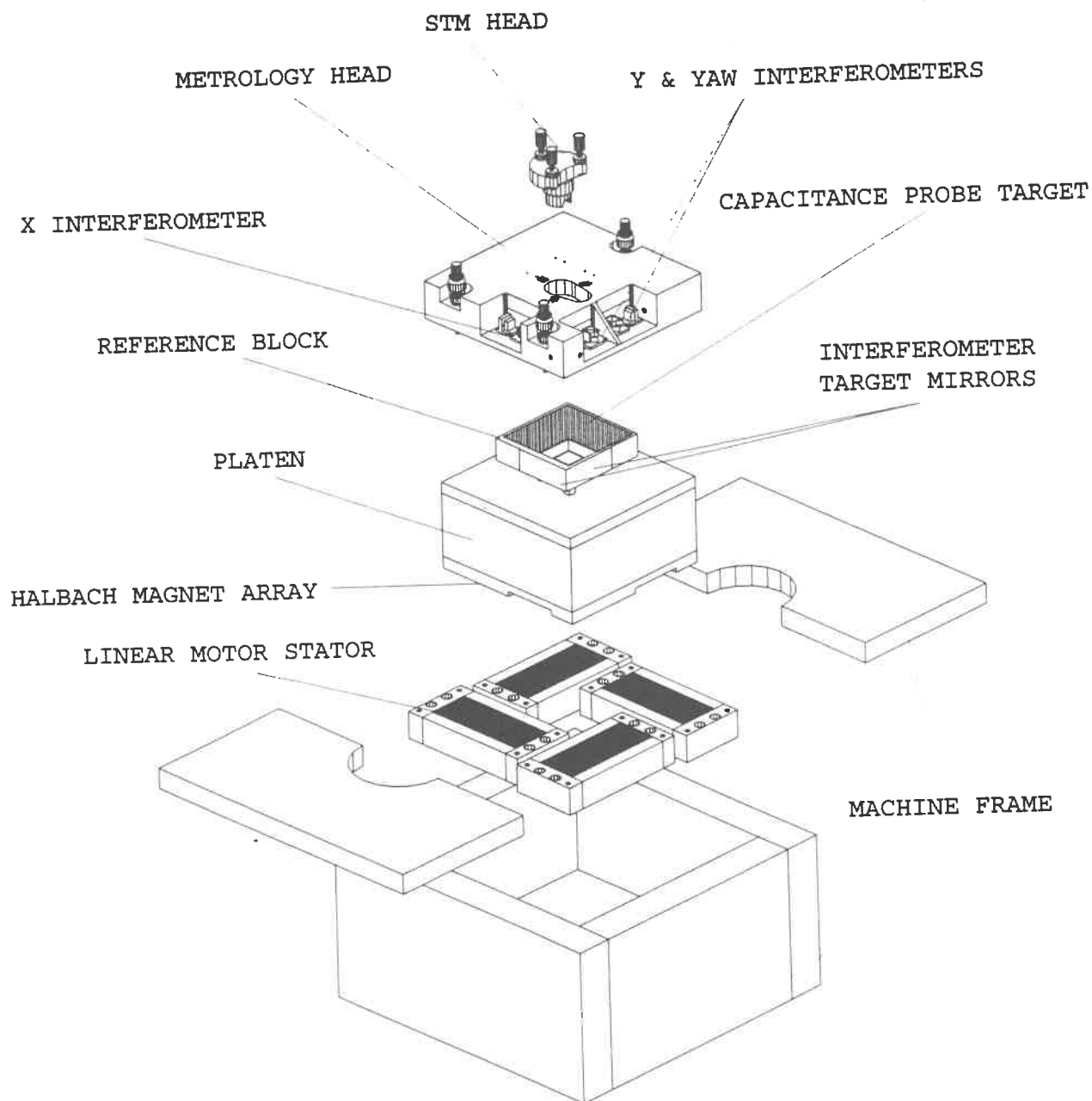


Figure 2: The LORS Stage (Exploded View)

Linear Motor Design

Four linear motors⁶, of the type shown in Figure 3, are used to suspend and servo the moving element (platen). The design procedure is outlined by Trumper et al⁷ where the design of a two phase linear motor is presented. For the sake of brevity, only the key features of the LORS Stage linear motors are presented here, the reader is referred to the reference for a detailed treatment of linear motor design.

Each of the linear motors consist of a stator and a permanent magnet array. The permanent magnet array consists of 16 NdFeB permanent magnets arranged in the Halbach configuration⁷. Positioned 500 μm directly beneath the magnet array is the stator (in

Figure 3 the gap between the magnet array and the stator is exaggerated for clarity). The stator consists of 66 coils each of which contains 10-turns of 18 gauge wire. Every sixth coil is connected in anti-series to make a six-phase linear motor. The motor phase resistance is 0.4Ω and the phase inductance is calculated to be approximately $30\ \mu\text{H}$. The linear motor shown is capable of exerting bidirectional vertical and horizontal forces F_v and F_h as indicated. This motor is designed for a gain of approximately 6 N/A with nominal operational forces of less than 1 N. The combined power dissipation of the four motors when the platen is suspended and driven with a lateral velocity of 1 mm/sec is calculated to be approximately 30 mW. A fixture has been built which will allow us to characterize the force-current-gap relationships for each of the linear motors. This fixture will also be used to measure unmodelled forces resulting from imperfect motor fabrication and assembly.

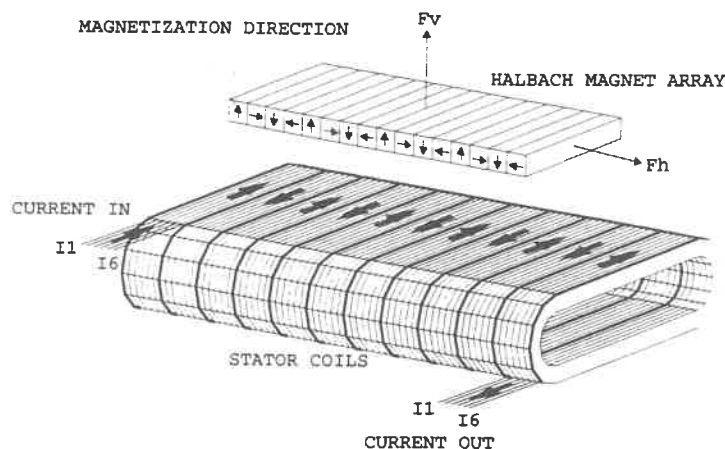


Figure 3: LORS Stage Linear Motor Configuration

Metrology

The LORS Stage metrology scheme is shown in Figure 4. The vertical position of the reference block is measured by the 3 capacitance probes. The capacitance probes are an integral part of the metrology head and are fabricated by depositing aluminum on the polished Zerodur surface. In Figure 4, the probe electrodes are shown in position over the reference block, although these electrodes are part of the metrology head. The overall probe diameter is 35 mm with a sensor diameter of 25 mm. The capacitance gage electronics are being provided by ADE Technologies⁹. The capacitance probes will be calibrated for a range of $100\ \mu\text{m}$ in air with a goal of 0.1 nm resolution inside a bandwidth of 10 Hz.

The lateral position of the reference block is measured by the 3 heterodyne interferometers. The 8 pass interferometer configuration was designed in collaboration with Zygo Corporation¹⁰. A Time-Interval-Analyzer (TIA) will be used to measure the relative phase of each of the measurement arm signals with respect to the reference arm signals. From this the lateral position and orientation of the reference block will be calculated with a resolution of 0.1 nm inside a bandwidth of 4 kHz.

Some Performance Predictions

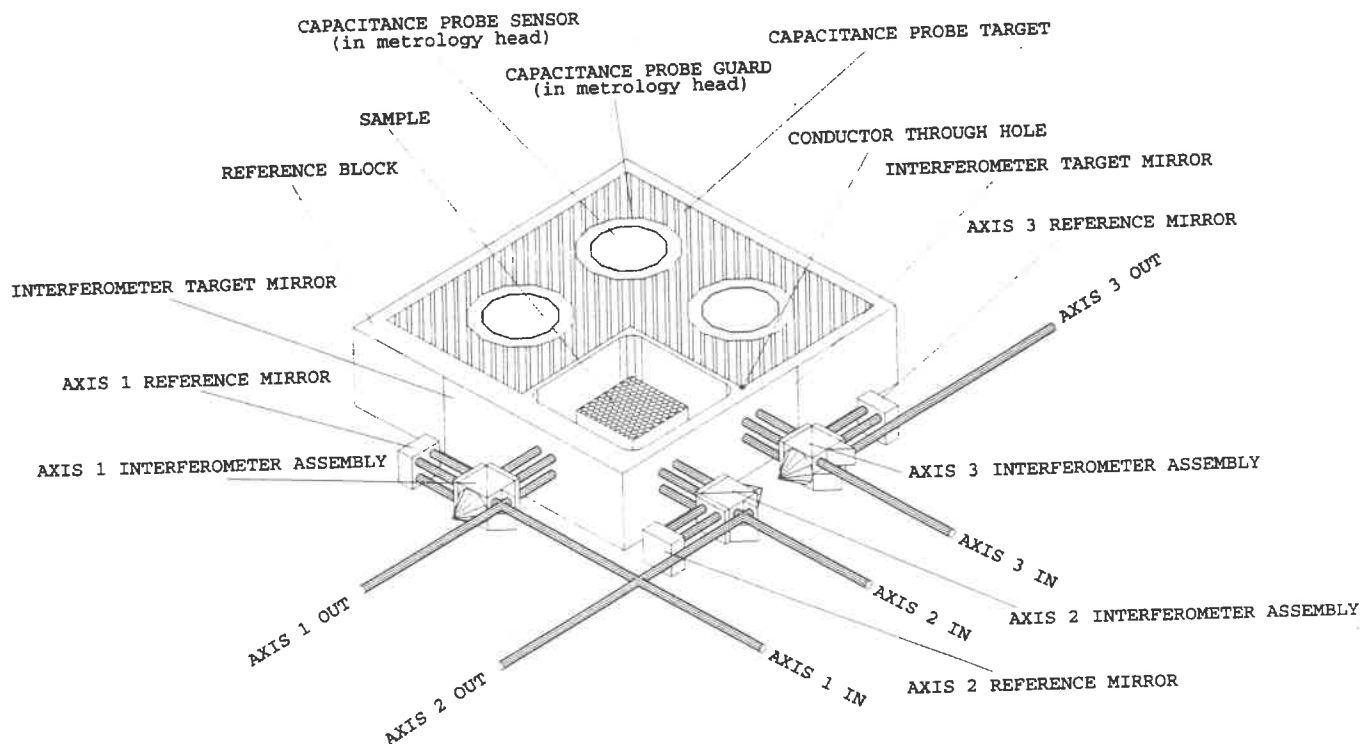


Figure 4: LORS Stage Metrology

The performance of the LORS stage will depend on sensor noise, disturbance forces, environmental influences, and systematic errors resulting from form errors of the reference block. Extensive modelling has been done in an effort to build a set of design tools which were utilized in the design phase of this project in an effort to achieve the performance criterion as stated in the introduction. These tools include a random noise model following the results of Ludwick et al^{11,12}, a heterodyne interferometer simulator^{13,14,15,16,17}, and a systematic error budget^{5,17,18}. Due to limited space in this extended abstract, only the results of these models are given herein.

The random noise model predicts a vertical position standard deviation of 0.024 nm and a horizontal standard deviation of 0.056 nm. The heterodyne interferometer simulator predicts, for the LORS stage interferometers, a temperature sensitivity of $1.2 \times 10^{-8} \text{m}/^{\circ}\text{C}$ in air and a pressure sensitivity of $2.5 \times 10^{-7} \text{m}/\text{psi}$. The simulator also predicts a 2.5 nm nonlinearity due to beam mixing. Note that the temperature and pressure errors were calculated for a maximum beam air-path imbalance of 12.5 mm. The systematic error budget predicts maximum lateral position errors of 4 nm and 7 nm and a maximum vertical position error of 10 nm.

The random noise predictions are conservative and we expect better performance than predicted. The temperature and pressure sensitivity will become less significant if we run the interferometers in a Helium filled constant pressure vessel which is currently planned. The systematic errors will be mapped in an effort to compensate for them and thereby lessen their influence.

Acknowledgments

This research is being funded by the National Science Foundation. The ADE Corporation

is providing the capacitive gaging electronics and Zygo Corporation is assisting with the metrology head and reference block fabrication.

References

- [1] Smith, S.T., and Chetwynd, D.G., *FOUNDATIONS OF ULTRAPRECISION MECHANISM DESIGN*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1994.
- [2] Personal communication with Dr. Stuart Smith.
- [3] Holmes, M.L., Trumper, D.L., and Hocken, R. J., "Atomic-scale precision motion control stage (The Angstrom Stage)," *CIRP Ann.*, vol. 44, no. 1, pp.455-460, 1995.
- [4] Holmes, M.L., and Trumper, D.L., "Magnetic/fluid-bearing stage for atomic-scale motion control (the Angstrom stage)," *Precision Eng.*, vol. 18, no. 1, pp. 38-49, Jan. 1996.
- [5] Slocum, A., *Precision Machine Design*, Prentice Hall, Englewoods Cliffs, NJ, 1992.
- [6] Trumper, D.L., "Magnetic Suspension Techniques for Precision Motion Control," Ph.D. Thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., Camb., Mass., Sept., 1990.
- [7] Trumper, D.L., Williams, M.E., and Nguyen, T.H., "Magnet Arrays for Synchronous Machines," in *Proc. IEEE IAS 28th Annu. Meet.*, Oct. 1993, pp.9-18.
- [8] Trumper, D.L., Kim, W.J., and Williams, M.E., "Design and Analysis Framework for Linear Permanent-Magnet Machines," in press, *IEEE Trans. Ind. Applic.*, vol. 32, no. 2, Mar/Apr 1996.
- [9] ADE Technologies, 77 Rowe Street, Newton, Mass. 02166.
- [10] Zygo Corporation, Laurel Brook Road, Middlefield, Conn. 06455.
- [11] Ludwick, S.J., and Trumper, D.L., "Noise Optimal Control in a Fluid-Floated Magnetic Positioner," in *ASME International Mechanical Engineering Congress and Exposition*, Atlanta, GA., Nov. 96.
- [12] Ludwick, S.J., Trumper, D.L., and Holmes, M.L., "Modeling and Control of a Six Degree of Freedom Magnetic/Fluidic Motion Control Stage," *IEEE Transactions on Control Systems Technology*, vol. 4, no. 5, Sept. 1996.
- [13] Howard, L.P., and Stone, J.A., "Computer modeling of interferometer errors using matrix methods," *Proceedings ASPE Annual Meeting*, pp. 143-146, 1995.
- [14] Personal communication with Dr. Lowell Howard at NIST.
- [15] Estler, W.T., "High-accuracy displacement interferometry in air," *Applied Optics*, Vol. 24, No. 6, 15, March 1985.
- [16] Bobroff, N., "Recent advances in displacement measuring interferometry," *Meas. Sci. Technol.*, 4, 907-926, 1993.
- [17] Personal communication with Dr. Robert Hocken.
- [18] Evans, C., Hocken, R., and Estler, T., "Self-Calibration: reversal, redundancy, error separation, and "absolute testing,"" *CIRP Ann.*, vol. 45, no. 2, 1996.

A Study on a High-Precision Linear Motion Table with Magnetic Bearing Suspension

Yoon Joung, Hyeong-Joon Ahn*, Owi-Serg Kim*, In-Bae Chang** and Dong-Chul Han***

Ultraprecision group, Production Engineering Center, Samsung Electronics Co., Suwon, Korea

*Graduate Student, Dept. of Mech. Design and Production Eng., Seoul National Univ.

** Dept. of Precision Mechanical Eng., Kangwon National Univ., Chuncheon, Korea

***Dept. of Mech. Design and Production Eng., Seoul National Univ., Seoul, Korea

Abstract

This paper describes a design, analyses and performances of a new type linear motion table with magnetic bearing suspension. Linear motion table with 250mm stroke in X axis is driven by linear motor and 5 d.o.f in other axes are actuated by magnetic bearing at arbitrary X position. The motor drives the table through its center of mass so that the driving forces may not cause any moment on the table which lead to Abbe errors on the table surface. The table consists of a platen and four magnetic bearing units and Each unit which consists of 3 electromagnets and 3 capacitance sensors affords control of forces by detecting the air gap between the table and guideway. Modeling of the 5 d.o.f system is carried out and compared with the experimental results in this paper. The performance test including positioning accuracy and resolution of the table system shows the feasibility of utilizing magnetic bearing suspension in high precision semiconductor manufacturing equipments.

1. Introduction

Magnetically levitated table system has lots of advantages when it is used in high precision machines like wafer steppers, IC measurement systems and scanned probe microscopy. A fully levitated stage requires no lubricants, which generates wear particles, and is well suited in high clean or vacuum environment. Also it has perfectly no friction and shows no stick-slip motion so that relatively high resolution can be obtained in comparison with other mechanical bearings. Many researches in this field are under way recently, and show good performances.[1-4]

This paper describes the design, analysis and performance of a new type 5 d.o.f linear motion table with magnetic bearing suspension. The table of 250mm stroke is driven by a linear motor and its position is detected by a linear scale. The table con-

sists of platen and four magnetic bearing units which are composed of electromagnets and capacitance sensors. Each magnetic bearing is controlled by displacement feedback signal from each capacitance sensor placed in line with electromagnets. One main axis of the table is driven by a linear motor in 250mm stroke, and other 5 axes are actuated by 6 magnetic bearings by controlling bias current in each electromagnet. The positioning accuracy and resolution of the table depends not on bearings but only on linear scale.

2. System Design

Table system. The typical system of interest here consists of guideways, supporting bed, linear motor with linear scale, and moving table with built-in gap sensors and electromagnets. Figure 1 shows the schematics of the linear motion table system.

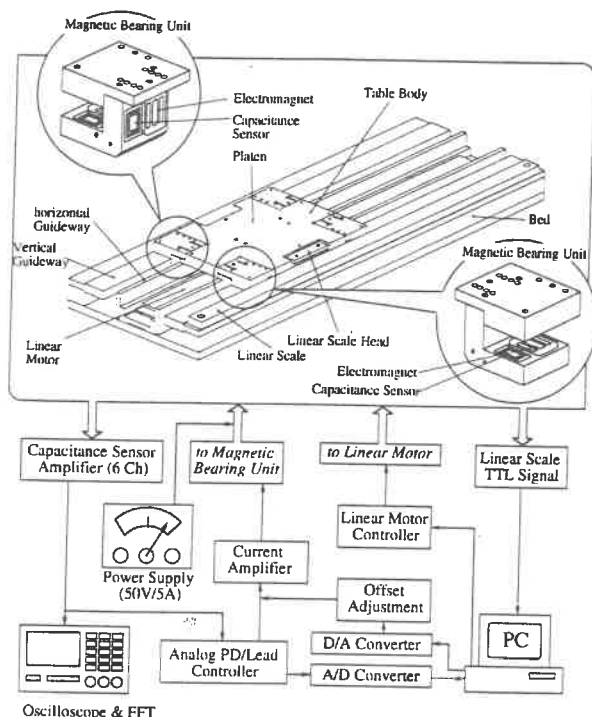


Fig. 1 Schematics of the linear motion system

The table consists of platen and four magnetic bearing units which can act as a suspension and 5 d.o.f actuators. Linear motor drives the table along the horizontal guideway through its center of mass according to its position feedback signal of the linear scale. In consequence, one d.o.f (long stroke) of the table is driven by linear motor and the other 5 d.o.f are actuated by each magnetic bearing.

Magnetic Bearing unit. Each magnetic bearing unit in figure 1 is composed of 3 electromagnets and 3 sensor plates. Two electromagnets placed opposite to each other in vertical axis act as a vertical bearing which support the platen by controlling the electromagnetic forces. Also one electromagnet placed in the side of this unit act as a horizontal bearing together with one in the other unit. Capacitance sensor plates are placed in line with each electromagnet to feedback the displacement signal of the table to the controller immediately.

Controller & Gap sensor. 6 Ch capacitance gap sensors are used to feedback the displacement signal

to the controller. The plates of the sensor with amplifiers are made in our laboratory, performance of which was proven in an early report[2]. To improve a linearization characteristic of the sensors, the signal of each sensor plate was differentially amplified so that two sensor plates become one gap sensor.

As shown in figure 1, each sensor signal is amplified and transmitted to the analog P/D controller which has offset adjustment, P/D, lead & phase compensation and filter circuits. The sensors placed close to each magnetic bearing allow us to use simple analog controller so that the table system can be made very simple and reliable.

Table. 1 Specifications of the table system

Table system	Table size	250 × 250mm
	Stroke	250mm(X axis) 0.1mm(Y,Z axis)
	Resolution	0.25 μ m
	Accuracy	< $\pm 1 \mu$ m
Electro magnet	Clearance	0.6 mm
	Core	0.35mm thickness Si steel 72 sheets
	Coil	ϕ 0.35 330 turns
	Pole area	320 mm ²
Gap sensor	Bandwidth	1.2 kHz
	Linearity	1%
	Linear range	± 0.05 mm
Linear motor	Resolution	0.1 μ m
	Stroke	250mm
	Peak force	30 lbs
Linear Scale	Acceleration	1 g
	Resolution	0.25 μ m
	Accuracy	$\pm 1 \mu$ m

Linear motor & Linear scale. To get 250mm stroke of the table, linear brushless D.C. motor is used which has simple structure and has no contact nature. Normag BLC04-32 motor with 30 lbs of peak force, has no attraction force in any other axis than driving axis so that the motor gives no disturbance force to the magnetic bearings. Heidenhein LS405 linear scale with 10 μ m of grating period shows good performance to use in position feedback system of less than 1 μ m order accuracy.

3. System Modeling

Figure 2 shows the coordinate system for 5 d.o.f modeling of the table.

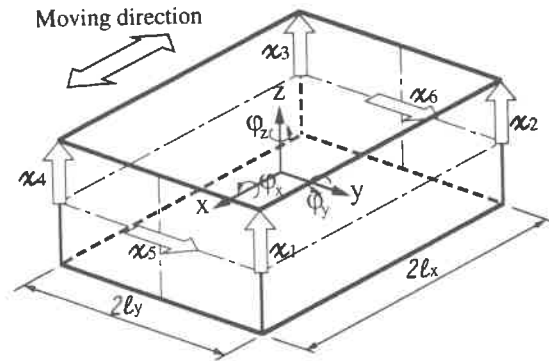


Fig.2 Coordinate system for 5 d.o.f modeling

The displacement of actuating point in Cartesian coordinate is $\vec{x} = [y \ z \ \varphi_x \ \varphi_y \ \varphi_z]^T$ and local displacements, which consist of four vertical and two horizontal direction, are $x_1 \sim x_6$, where electromagnets are placed. Once a position vector $\vec{R}_i$ and a direction vector $\vec{N}_i$ of each actuator are defined, local force, (f) , generated by each bearing and external force, $\vec{F}$, on Cartesian coordinate are related as follows;

$$\vec{N}_a \cdot (f) = \vec{F} \quad (1)$$

where actuator characteristic matrix $\vec{N}_a$ is defined as,

$$\vec{N}_a = (\vec{Q}_1, \vec{Q}_2, \dots, \vec{Q}_n) \quad (2)$$

$$\vec{Q}_i = [\vec{N}_i^T, (\vec{N}_i \times \vec{R}_i)^T]^T$$

The sensor signal s_i can be defined with sensor characteristic matrix $\vec{N}_s^T$ as follows,

$$(s) = (s_1, s_2, \dots, s_n)^T = \vec{N}_s^T \cdot \vec{x} \quad (3)$$

$$\vec{N}_s^T = (\vec{N}_1^T, (\vec{N}_1 \times \vec{R}_1)^T) \quad (4)$$

$\vec{N}_a, \vec{N}_s^T$ are function of the distance of electromagnets l_x, l_y .

Governing equation of the system is as follows;

$$M\ddot{\vec{x}} + K\vec{x} = \vec{F} \quad (5)$$

$$M = \text{diag}[m \ m \ J_x \ J_y \ J_z]$$

$$K = \text{diag}[2k_{yy} \ 4k_{zz} \ 4k_{zz}l_y^2 \ 4k_{zz}l_x^2 \ 4k_{yy}l_x^2]$$

where m is mass of the table, J_x, J_y, J_z are mass moment of inertia with respect to axes x, y and z . The position stiffness at each direction, k_{yy} and k_{zz} , is the force variation due to the table displacement at nominal table position.

Finally figure 3 shows the block diagram for system modeling.

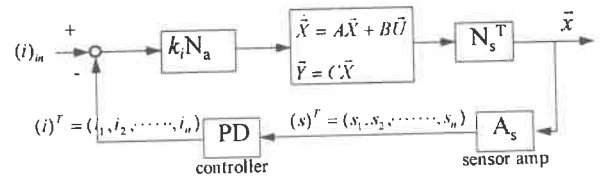


Fig.3 Block diagram for system modeling

where the state variables are

$$\vec{X} = [y \ z \ \varphi_x \ \varphi_y \ \varphi_z \ \dot{y} \ \dot{z} \ \dot{\varphi}_x \ \dot{\varphi}_y \ \dot{\varphi}_z]^T$$

and k_i is an actuator gain which is a function of bias current.

4. Experimental Results

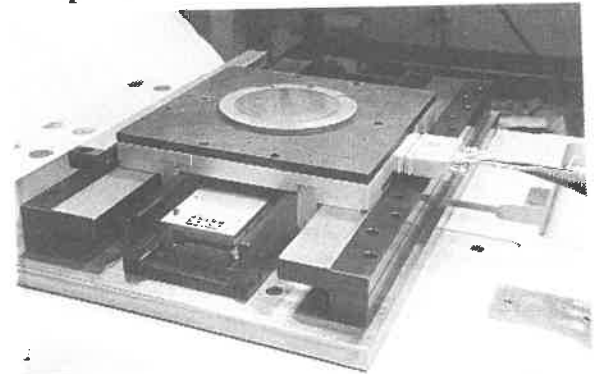


Fig. 4 General view of the table system

Figure 4 shows the general view of the linear motion table system.

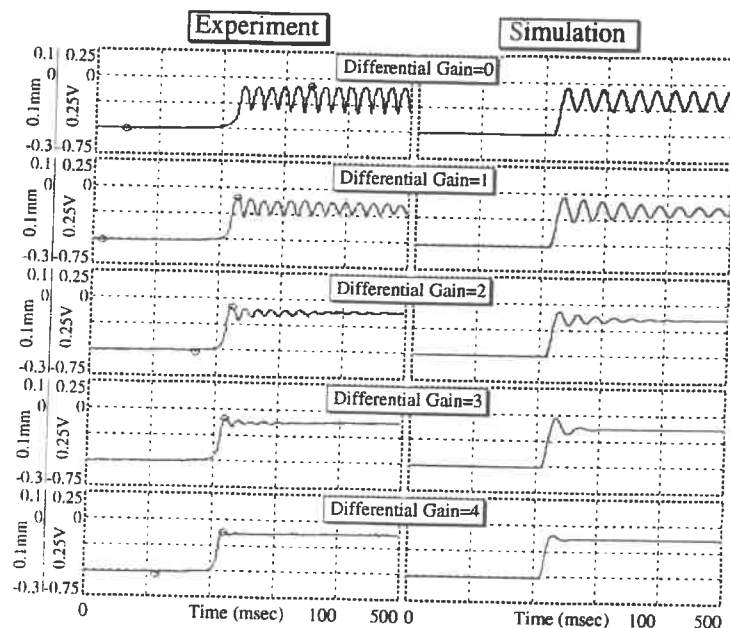


Fig. 5 Step response of the table system

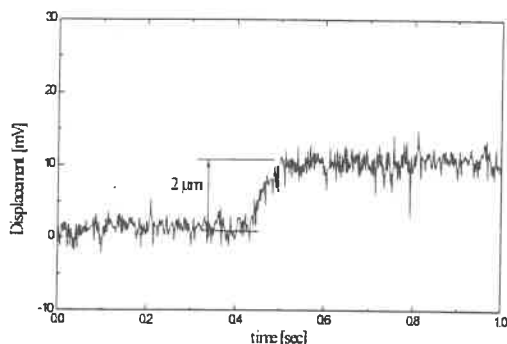


Fig. 6 $2\mu\text{m}$ step in Z-axis

The result of step response of the system by experiment and simulation is shown in figure 5. The validity of analytical results are confirmed by comparison with experimental results. The damping characteristics could be improved by increasing the differential gain of the controller, however, the system becomes unstable when the gain is more larger than 5. Figure 6 shows $2\mu\text{m}$ step in Z-axis, perpendicular to the main axis, by actuating electromagnet with DAC of 12 bits resolution. Noise component in figure 6 is ground noise which limits the position resolution of the bearings.

5. Conclusions

Linear motion table with magnetic bearing sus-

pension and built-in capacitance gap sensor has been developed. Linear motor and linear scale of $0.25\mu\text{m}$ resolution are adopted to get 250 mm stroke in X-axis and other 5 axes are controlled by each magnetic bearing in 0.1mm stroke. Analytical results of 5 d.o.f modeling and experimental result show good agreement.

Acknowledgements

This work was funded by Samsung Electronics Co. and supported in part by Turbo & Power Machinery Research Center and Institute of Advanced Machinery & Design in Seoul National University.

References

- [1] M.E.Williams, D.L.Trumper. "Precision magnetic bearing six degree of freedom stage", *3rd Int'l Symp. on Magnetic Suspension Tech.*, July, 1995, pp.453.
- [2] Y.Joung, S.C.Jung, I.B.Chang, D.C.Han, "An experimental study on the high-precision linear motion table with magnetic bearing suspension", *Proceedings of ASPE 10th Annual Meeting*, Oct., 1995, pp.328.
- [3] D.Eisenhaure, A.Slocum, A.R.Hockney, "Magnetic Bearings for Precision Linear Slides", *1st Int. Symposium on Magnetic Bearings*, 1988, pp. 67.
- [4] S.Shingu, "Design Method of Magnetically Suspended Positioning Mechanism", *Journal of JSME*(in Japanese), Vol.55, No.518, 1989, pp.2559.
- [5] I.B.Chang, "A study on the performance improvement of a magnetic bearing system using built-in capacitive type transducer", Ph.D. diss., Seoul Nat'l Univ., 1994, Seoul. Korea

FABRICATION AND THERMAL ANALYSIS OF A GAS-LIQUID PHASE-CHANGE MICRO-ACTUATOR

Shigeo KATO

Nippon Institute of Technology,
Miyashiro, Saitama 345, Japan

1. Introduction

When a liquid is vaporized, the volume of the phase-changed gas is more than 200 times that of the liquid. Gas-liquid phase-change actuators are thermal engines that utilize the large change in volume that occurs during a phase-change. These actuators are expected to be applied as the micro-actuators used in micro-robots operating in very small pipings or the human-body, because conventional electric power, laser beams or microwaves can be used as the source of vaporizing energy. So far, these actuators have been applied as micro-pumps for very small flow rates. Mizoguchi et al. fabricated a micro-pump whose flow rate was 30 nl per cycle [1]. We have fabricated a prototype of a gas-liquid phase-change micro-actuator that supplies an actuating force of 3 N to micro-robots sufficient to operate in the human-body. This paper presents a fabrication and thermal analysis of the gas-liquid phase-change micro-actuator.

2. Characteristics of the fabricated actuator

A cross section of the fabricated micro-actuator is shown in Fig. 1. The casing including the flange is made of aluminum. It has a liquid pool with a volume of 79 μ l. The volume of 79 μ l is almost the same as that of two or three drops from an eye-dropper. The operating fluid is R-113, because its vaporizing temperature, though higher than the human body, is still only 47.56 $^{\circ}$ C and its latent heat for vaporization is only 1/16 that of water. Consequently, the energy needed for vaporization is very low. The operating liquid is heated by a heater 0.2-mm thick and 136-mm long made of coiled constantan wire with an electrical resistance of 2.1 W. The force of volume-change is received by a 0.2-mm thick artificial rubber-membrane, then transferred to the stroke of a small piston.

An electrical power of 2.1 W is supplied with in 1 second to the heater and it is cut after 3 seconds. This operation is repeated. The maximum displacement of the piston gradually increases. After 60 seconds, the maximum displacement reaches 1 mm and the minimum displacement reaches 0.44 mm. This motion is shown in Fig. 2 by the solid line.

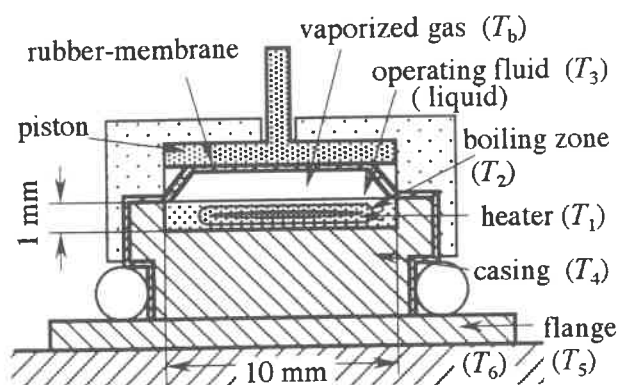


Fig. 1 A gas-liquid phase-change micro-actuator

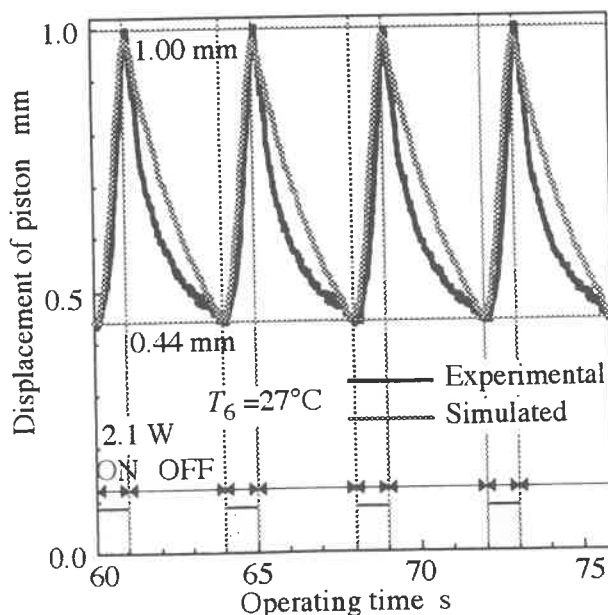


Fig. 2 Continuous stroke motion of the piston

Thus, this actuator is confirmed to have a reciprocating motion with a 0.5-mm stroke at an average power of 0.525 W. In order to make a mathematical thermal model of the actuator, 1-stroke motion was examined in an experiment. If the heater is heated by 2.1 W of electricity, the piston starts to move up 0.41 seconds after power is supplied at an ambient temperature of 26°C and its displacement reaches 1 mm 2.2 seconds after power is supplied.

By switching off the power, the piston begins to move downward after 0.2 seconds delay. Its displacement is 0.4 mm after 5 seconds, and 0.15 mm 10 seconds after the power switched on. This motion is indicated by the solid line in Fig. 3.

If the power decreases to 0.5 W from 2.1 W, the stroke of the piston freezes. The performance of the actuator is considered to be thermally critical. In this case, the ambient temperature (T_6) is 25 °C, the temperature of the flange is 26.35 °C, and the temperature of the operating liquid is 37 °C.

3. Mathematical model of the actuator

In order to simulate the actuator, we needed to create a mathematical thermal model of the actuator. A model of the actuator is shown in Fig. 4. The temperature distribution is very complicated. However, we simply considered five thermal elements receiving the heat and their temperature changes. These are (1) the heater, (2) the operating liquid near the heater (boiling zone), (3) the operating liquid near the casing, (4) the actuator-casing, and (5) the flange for dissipating the heat. If we select only one liquid for the operating fluid, we cannot explain why vaporization starts 0.41 seconds after power is supplied. The heat capacity, the temperature, and the heat transfer area of the element- i are $C_i(J/°C)$, $T_i(°C)$, and $A_i(m^2)$, respectively. The heat flow rate and the heat transfer coefficient from element- i to element- $(i+1)$ are $q_i(W)$, and $h_i(W/m^2°C)$, respectively. The heat flow rate to the heater is $q_H(=2.1 W)$. The heat flow rate for vaporizing is $q_v(W)$. The heat flow rate when the vaporized gas is re-liquefied and the liquefying area are $q_L(W)$ and $A_L(m^2)$, respectively. Simultaneous differential equations of the thermal model are solved by the Runge-Kutta method.

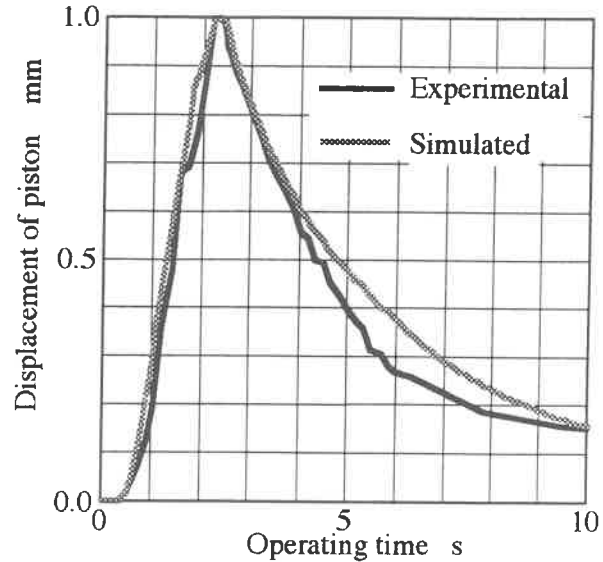


Fig. 3 Displacement of piston versus the operating time

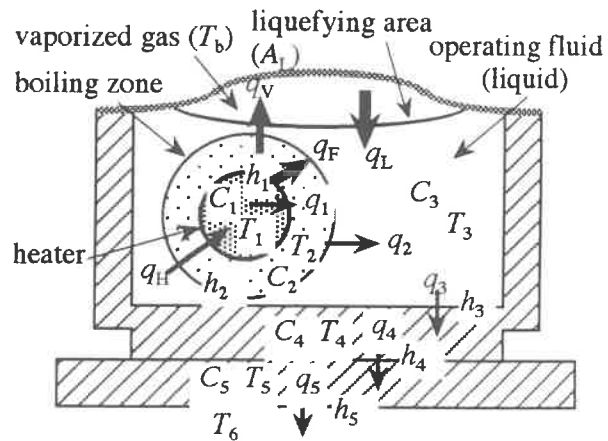


Fig. 4 Thermal model of the actuator

3.1 Under power supply (from start to vaporization)

When power is supplied to the heater, the operating liquid near the heater quickly vaporizes. For the period from the start till vaporization occurs, the heat balances of the five thermal elements are represented by next five differential equations. t represents the operating time.

$$C_1 \frac{dT_1}{dt} = q_H - q_1 = q_H - A_1 h_1 (T_1 - T_2) \quad (1)$$

$$C_2 \frac{dT_2}{dt} = q_1 - q_2 = A_1 h_1 (T_1 - T_2) - A_2 h_2 (T_2 - T_3) \quad (2)$$

$$C_3 \frac{dT_3}{dt} = q_2 - q_3 = A_2 h_2 (T_2 - T_3) - A_3 h_3 (T_3 - T_4) \quad (3)$$

$$C_4 \frac{dT_4}{dt} = q_3 - q_4 = A_3 h_3 (T_3 - T_4) - A_4 h_4 (T_4 - T_5) \quad (4)$$

$$C_5 \frac{dT_5}{dt} = q_4 - q_5 = A_4 h_4 (T_4 - T_5) - A_5 h_5 (T_5 - T_6) \quad (5)$$

3.2 Under power supply (boiling zone is partial in the pool)

When the temperature of the operating liquid near the heater reaches the boiling temperature ($T_b = 47.56^\circ\text{C}$), vaporization starts. When vaporization starts, the relating heat transfer coefficients h_1 and h_2 are assumed to change to h_{1v} and h_{2v} , respectively.

$$C_1 \frac{dT_1}{dt} = q_H - q_1 = q_H - A_1 h_{1v} (T_1 - T_b) \quad (6)$$

$$C_2 \frac{dT_2}{dt} = 0 \quad (7)$$

$$C_3 \frac{dT_3}{dt} = q_2 - q_3 + q_L = (A_2 h_{2v} + A_L h_L) (T_b - T_3) - A_3 h_3 (T_3 - T_4) \quad (8)$$

$$C_4 \frac{dT_4}{dt} = q_3 - q_4 = A_3 h_3 (T_3 - T_4) - A_4 h_4 (T_4 - T_5) \quad (9)$$

$$C_5 \frac{dT_5}{dt} = q_4 - q_5 = A_4 h_4 (T_4 - T_5) - A_5 h_5 (T_5 - T_6) \quad (10)$$

$$q_1 - q_2 = A_1 h_{1v} (T_1 - T_2) - A_2 h_{2v} (T_2 - T_3) = q_v + q_F \quad (11)$$

The zero in Eq. (7) shows that temperature T_2 maintains the boiling temperature, and $(q_1 - q_2)$ is obviously not equal to zero. The term q_F in Eq. (11) is the heat flow rate that is used to spread the vaporizing zone (the operating liquid near the heater). Supposing that β is a constant less than 1, it is assumed that terms q_v and q_F can be represented as follows.

$$q_v = \beta (q_1 - q_2) \quad (12)$$

$$q_F = (1 - \beta) (q_1 - q_2) \quad (13)$$

Increasing of the vaporizing zone is represented as

$$(T_b - T_3) \frac{dC_2}{dt} = q_F = (1 - \beta) (q_1 - q_2) \quad (14)$$

The mass per unit of time of the vaporizing fluid (kg/s) is represented as

$$m_v = \frac{q_v - q_L}{H_b} \quad (15)$$

where H_b is the latent heat for vaporization (kJ/kg $^\circ\text{C}$) of the operating fluid.

Table 1 Simulation data

Heat Transfer Area	$A_1=85.4$ $A_2=85.4$ $A_3=109.9$ $A_4=95$ $A_5=530$ $A_{L1}=78.5$ $\times 10^{-6} \text{ m}^2$
Heat Capacity	$C_1=0.0156$ $C_2=0.000128$ $C_3=-0.111$ $C_4=1.252$ $C_5=2.582$ $\text{J/}^\circ\text{C}$
Heat Transfer Coefficient	$h_1=560$ $h_{1v}=1120$ $h_2=549$ $h_{2v}=1098$ $h_3=432$ $h_{3v}=880$ $h_4=30000$ $h_5=699$ $h_L=80$ $\text{W/m}^2 \text{ }^\circ\text{C}$
Ratio of Heat Flow Rate for Vaporization	$\beta=0.103$

3.3 Terminated power supply (no boiling zone)

$$C_1 \frac{dT_1}{dt} = -q_1 = -A_1 h_1 (T_1 - T_3) \quad (16)$$

$$C_2 \frac{dT_2}{dt} = 0 \quad (17)$$

$$C_3 \frac{dT_3}{dt} = q_1 - q_3 + q_L = A_1 h_1 (T_1 - T_3) - A_3 h_3 (T_3 - T_4) + A_L h_L (T_b - T_3) \quad (18)$$

$$C_4 \frac{dT_4}{dt} = q_3 - q_4 = A_3 h_3 (T_3 - T_4) - A_4 h_4 (T_4 - T_5) \quad (19)$$

$$C_5 \frac{dT_5}{dt} = q_4 - q_5 = A_4 h_4 (T_4 - T_5) - A_5 h_5 (T_5 - T_6) \quad (20)$$

$$q_L = A_L h_L (T_b - T_3) \quad (21)$$

$$m_v = -\frac{q_L}{H_b} \quad (22)$$

Eq. (17) shows that C_2 is equal to 0.

4. Supposition of the heat transfer coefficients

Using the heat transfer areas shown in Table 1 and critical power of 0.5 W, some heat transfer coefficients have been supposed. We consider that T_2 is at a boiling temperature of 47.56°C . Under these conditions, h_2 is 549, h_3 is 432, h_4 is 3×10^4 , and h_5 is 699 ($\text{W/m}^2 \text{ }^\circ\text{C}$).

The heat transfer coefficient h_1 is estimated from the start-up time of the piston. The initial

thickness of the fluid near the heater is considered to be 1, 5 or 10 μm . The relationship between the heat transfer coefficient h_1 and the start-up time of the piston is simulated. Since we know $h_1=572$ ($\text{W/m}^2\text{°C}$) from the calculation and the start-up time of the piston is 0.41 s, the initial thickness of the fluid near the heater may be about 1 μm . We decided to use $h_1=560$ ($\text{W/m}^2\text{°C}$) and 1 μm for the initial thickness of the fluid near the heater for simplicity.

If the heat transfer coefficient h_3 increases, heat transfer to the casing increases and vaporization decreases. Since the displacement of the piston reaches 1 mm at $t=2.2$ s, we know that the heat transfer coefficient h_{3v} is equal to 880 ($\text{W/m}^2\text{°C}$) by simulation. This value is about two times $h_3=432$ ($\text{W/m}^2\text{°C}$) which is the heat transfer coefficient prior to boiling. Consequently, we supposed that h_{1v} and h_{2v} which are heat transfer coefficients may be two times the heat transfer coefficients prior to boiling, i.e., $h_{1v}=1120$, $h_{2v}=1098$ ($\text{W/m}^2\text{°C}$), respectively.

The heat transfer area for liquefying A_L is an area of 10-mm diameter ($A_{L1}=78.5\times10^{-6}\text{m}^2$) if the displacement is as large as $H=1$ mm. However, if the displacement of the piston decreases, the area where the gas and liquid come in contact with each other may decrease due to the capillary phenomenon. If the displacement is zero, then area A_L may be zero. Consequently, we assume that area A_L is related to displacement H and is represented as

$$A_L=A_{L1}\left(\frac{H}{H_1}\right)^2 \quad (23)$$

where H_1 shows a displacement of 1-mm.

The heat transfer coefficient for liquefying h_L is assumed to be $h_L=80$ ($\text{W/m}^2\text{°C}$), because this value enables the displacement of the piston to reach $H=0.15$ mm 10 s after power supply is supplied.

5. Simulation of the displacement of the piston

We simulated the displacement of the actuator using the values shown in Table 1. We used a value of $\beta=0.103$, because this value enables the displacement of the piston to reach $H=1$ mm 2.2 s after energy is supplied.

The relationship between volume V (μl) of the vaporized gas and displacement H (mm) is represented as

$$V=69 H \quad (24)$$

The relationship between the pressure in the rubber membrane P (MPa) and the displacement H (mm) is represented as

$$P=0.101 + 0.04H \quad (25)$$

The pressure and the volume are represented by the equation of state for the ideal gas as

$$PV=M_v R T_g \quad (26)$$

where M_v and R are the mass of the gas in the rubber membrane and the gas constant of the operating fluid, respectively.

Because the molecular weight of R-113 is 187.376, from Eqs. (24), (25), and (26), the relationship between the displacement H and the mass of the gas M_v (kg) is represented as

$$H(0.101 + 0.04 H) = 149050 M_v \quad (27)$$

If we know the mass of the gas, we obtain the displacement of the piston by Eq. (27).

The results of the simulation for one stroke of the piston are shown in Fig. 3 by the dotted line. We see a difference of about 0.1 mm at $t=6$ s. However, we know that although the simulation coincides with the experiment there is a small difference. The simulation of the continuous reciprocating movement of the piston is shown in Fig. 2 by the dotted line. It is confirmed that continuous reciprocating movement of the piston with a 0.5-mm stroke can be simulated.

6. Conclusions

(1) We fabricated a gas-liquid phase-change micro-actuator that is 16 mm in diameter and 10 mm thick, and that has a liquid pool with a volume of 79 μl . The operating fluid was R-113. This actuator generated a force of 3N.

(2) By dividing the operating fluid into two thermal elements, we could make a simple mathematical model for the gas-liquid phase-change micro-actuator.

(3) By experiment and simulation, the actuator was confirmed to reciprocate continuously with a 0.5-mm stroke at an average input power of 0.525 W.

Reference

- [1] H. Mizoguchi, M. Ando, T. Mizuno, T. Takagi, and N. Nakajima: Proc. of Micro Electro Mechanical Systems '92 pp. 31-36.

LOAD-DEPENDENT BEHAVIOUR IN PTFE-GLASS SLIDEWAYS

A. F. Monteiro and D. G. Chetwynd

Centre for Nanotechnology and Microengineering
Department of Engineering,
University of Warwick, Coventry CV4 7AL UK

Introduction. There has been much recent interest in using slideways for nanometre-precision linear translation [1,2]. Typically, these are constructed with an optically polished glass prismatic surface carrying thin-film PTFE bearing pads. Compared to elastic deformation mechanisms these offer the possibility of much larger ranges and lower parasitic tilts in yaw, roll and pitch. However, one characteristic of sliding motion, particularly significant at the nanometre level, is a hysteretic (lost motion) behaviour, documented in theoretical modelling and laboratory tests [3,4]. This sliding hysteresis (typically of the order of 100 nm) is a result of the stresses and strains present at the contact points between the moving carriage feet and the flat slideway surface and theory suggests that its magnitude is strongly dependent on the normal load at the contact points.

Reducing the hysteresis is important if there is a need to work open-loop to high levels of accuracy, but even when feedback is used a much flatter hysteresis cycle helps make the control system reliable and easier to design. Thus an experimental investigation has been carried out into how in practice reducing the normal force at the bearing pads of an ultra-precision slideway carriage will alter the sliding hysteretical curve.

Modelling. Each bearing pad of the slideway behaves as a Hertzian contact with a thin film entrapped between harder sphere and flat. Theoretical models of Hertzian contacts under transverse load were developed as early as 1953 by Mindlin and Deresiewicz [3], by making some plausible assumptions about the distribution of regions of microslip and of elastic contact, relaxing the solutions to the ensuing stress distributions in order to smooth discontinuities and then assuming that in the microslip regions the horizontal traction is given by the conventional limiting friction condition of the local normal load. They considered only the contact of spheres of identical material in the initial model, but given the nature of the other approximations, extending it to our more complex interface does not appear an extreme step.

Hertzian analysis shows that pressing a smooth sphere of radius R onto a smooth flat surface by a stationary perpendicular force P , causes a circular contact zone of radius a , given by

$$a = \left(\frac{3PR}{4E^*} \right)^{1/3} \quad (1)$$

where

$$\frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \quad (2)$$

and ν and E are the Poisson's ratio and elastic modulus of the two materials respectively. Applying a tangential force T to the sphere, would cause an infinite stress around the edge of the circular contact region unless microslip started there and as T increases, an annular region of slip grows towards the centre of the contact area with an inner radius of slip, c , given by

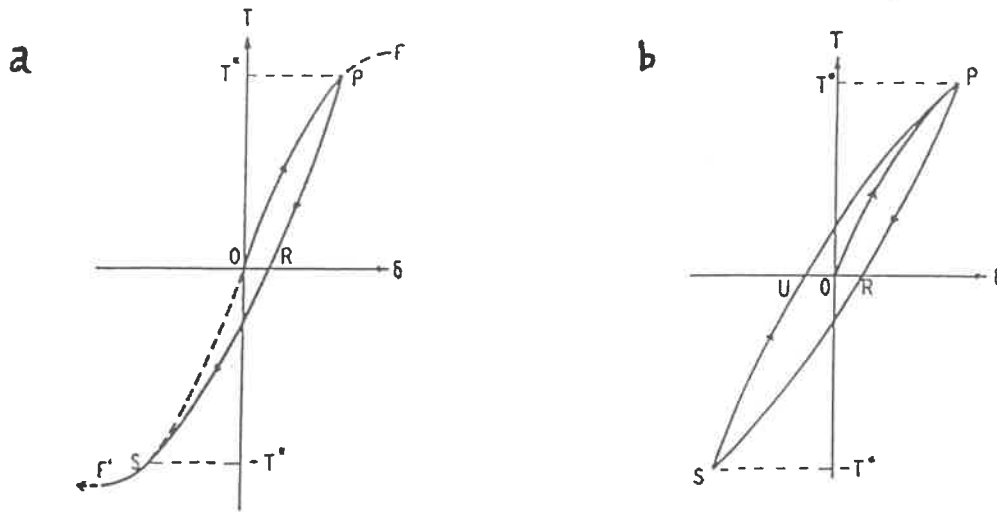


Figure 1: Hysteretic transverse displacement of sphere under microslip under (a) initial load and unload and (b) continued cyclical loading.

$$c = a \left(1 - \frac{T}{\mu P} \right)^{1/3} \quad (3)$$

where μ is the coefficient of friction. When T reaches μP , the inner radius reduces to a point and sliding occurs. For $T < \mu P$, Mindlin and Deresiewicz estimate the lateral displacement from elastic deflections of the non-slip region as

$$\delta_r = \frac{3(2-\nu)\mu P}{8Ga} \left(1 - \frac{c^2}{a^2} \right) = \frac{3(2-\nu)\mu P}{8Ga} \left[1 - \left(1 - \frac{T}{\mu P} \right)^{2/3} \right] \quad (4)$$

where G is the shear modulus (taken by us to be that of the interface material).

If we load initially to some value T^* and then decrease the load towards zero and into the opposite direction, similar analysis shows that there will be an offset in the displacement, figure 1a, that leads to a classic hysteresis curve if the cycle is repeated, figure 1b. We define the hysteresis (H) of this cycle as the maximum difference in the respective positions of the bearing at the same nominal point for the two directions of travel (twice the length OR in figure 1a) given by

$$H = 2 \times \frac{3(2-\nu)\mu P}{8Ga} \left[2 \left(1 - \frac{T^*}{2\mu P} \right)^{2/3} - \left(1 - \frac{T^*}{\mu P} \right)^{2/3} - 1 \right] \quad (5)$$

The cyclic lost motion from these elastic effects if gross sliding is allowed is found by setting T^* to μP , the limiting condition. As the radius of the bearing remains basically fixed for a given system, it can be absorbed, along with elastic properties, into a constant K , so

$$H = K\mu P^{2/3} \quad (6)$$

Experimental Method. Tests of sliding-direction hysteresis were carried out on a mechanism specially designed and built to study kinematic techniques for straight line motion with sub-nanometre and sub-milli-arc-second error motions [4], figure 2. It is made almost completely of Zerodur ultra-low-expansion glass ceramic, with sintered PTFE-bronze-lead bearing pads, spherically radiused and carrying a thin (generally sub-micrometre) film of PTFE across their surface. A 45° half angle optically flat prism is connected to a base plate forming the slideway

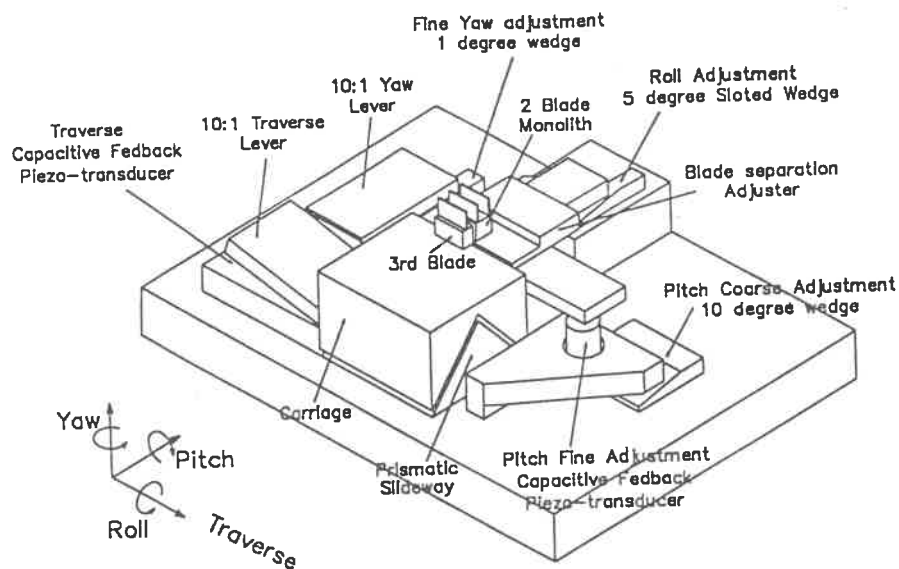


Figure 2: The super-precision kinematic slideway system [4].

assembly. A carriage block with a prismatic V cut from its underside rests on the slideway supported by five PTFE pads, three on one side of the prism and two on the other to form a kinematic arrangement with one degree of freedom. The carriage is driven by a piezoelectric actuator employing local capacitive feedback to provide linear motion at the nanometre level. This motion is then further reduced by a 10/1 reduction lever which pushes the carriage forward. The thrust-line of this lever is aligned closely along the carriage friction centre line. The carriage is pre-loaded against the driver by a simple helical spring, and its traverse is measured using a laser interferometer system, where a spot beam is bounced off a small mirror attached to the carriage. The stage was mounted on a granite table in a room temperature controlled to about ± 1 K and held at approximately 40% relative humidity. The stage was enclosed by a Perspex cover to reduce air currents and thermal drift around the interferometer light paths. The input and output interferometer beams go through the clear plastic enclosure while all uncompensated (i.e. not common path) beams remain inside. The Hewlett Packard interferometer has an ideal resolution of 10 nm, estimated to be reduced to a useful resolution of 20 nm under these operating conditions.

The system includes a weight reduction mechanism for the moving carriage, figure 3. Supported by a metrology frame, a non-rotating-spindle micrometer holds a yoke suspending three soft springs which are attached to three spring posts on the moving carriage. Raising the micrometer stretches the suspension springs which then support a proportion of the carriage weight, the rest of which is balanced by the normal force on the contact bearings. This force can be reduced to very low values while ensuring that the slideway remains in proper kinematic contact. While unsuitable for most practical bearings, this geometry allows very soft springing that makes negligible any thermally related effects in the lifting system and ensures that there is no detectable change in the lift as the carriage travels over the range of motion here being used. The springs were selected such that an easily measured extension of 5 mm was required to completely lift the carriage while keeping well within their linear (Hooke's law) mode of behaviour. The micrometer adjustment then gave a load adjustment resolution much higher than required and overall linearity can be assumed based just on the reading of the micrometer. The experimental

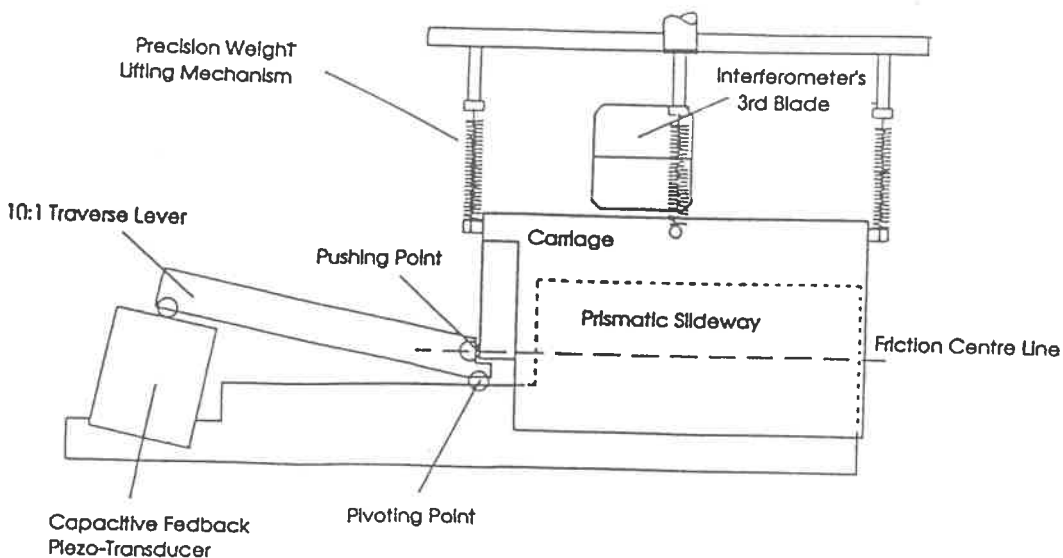


Figure 3: Details of the linear drive and weight reduction system.

investigation depends on the ability to change the load in known proportions, which is ensured by the linearity of the arrangement, and there is no special benefit in having a very precise knowledge of the load in absolute terms. Thus the force calibration consisted simply of weighing the carriage and then measuring the spring extension required to cause lift-off.

The carriage was driven back and forth through nominally ± 500 nm for two complete cycles while recording both the drive actuator displacement (from its internal capacitive gauge) and the resultant motion of the carriage (detected by the interferometer). The actuator control box was adjusted manually to give nominal positions and then the laser reading noted after a few seconds delay to ensure that the system had stabilized. This sequence was repeated for different settings of the micrometer head on the weight reduction mechanism, starting at a bearing load of around 0.5 N (9.5 mm at the head) and increasing it in steps of 0.5 N to 5.5 N (4.5 mm). Note that this procedure depends on the moderate stiffness of the coupling between drive and stage to relax sufficiently for the stage to follow the pads rather than the drive output.

Results and discussion. Figure 4 shows example cyclical plots using loads of 0.5 N and 5.5 N. Each cycle was performed twice to improve reliability and get a measure of the drift present in the interferometer system. The thermal drift over times of one or two minutes (attributable almost entirely to the interferometer) was roughly linear (experimental observation) and this first order effect can then be subtracted from the raw data (such as shown in figure 4) to obtain H . It was considered more practical to subtract modest drift from the measurements than to allow very long settling periods at each set-up. Observed drifts of up to half the width of the hysteresis loop in any measurement cycle were accepted for this procedure. The hysteresis is then the maximum separation of the higher and lower parts of each corrected cycle, which in practice always occurred at the mid-point of the cycle. It is plotted against the normal loading on the bearings in figure 5. The uncertainty in these readings is estimated to be around 20 nm standard deviation.

Experimentally, the load P is well represented by the displacement, d , of the micrometer supporting the weight-reduction springs. Taking the lowest micrometer reading at which the stage floats above the slideway with no normal loading as d_0 , equation 6 can be re-written as

$$H = K_2(d - d_0)^{2/3} \quad (7)$$

where K_2 is a constant incorporating K , μ and the combined stiffness of the three springs to the power of $2/3$. This model can be compared to the data by estimating best-fit values for K_2 and d_0 but simple procedures for doing this are very sensitive to errors in d_0 at low loads. However there is little sensitivity to the offset at higher loads and so iterative procedures converge rapidly. By such an approach, values of $d_0 = 9.95$ mm and $K_2 = -73.4 \times 10^{-7} \text{ m}^{1/3}$ were found to be a good empirical fit to the data. The plot of the experimental data, figure 5, includes this theoretical fit and also shows the residual errors of the data from the line. The quality of the agreement strongly suggests that the hysteretic behaviour of the complex combination of materials in our bearings is not functionally different from the simple, single material case considered in [3] and above.

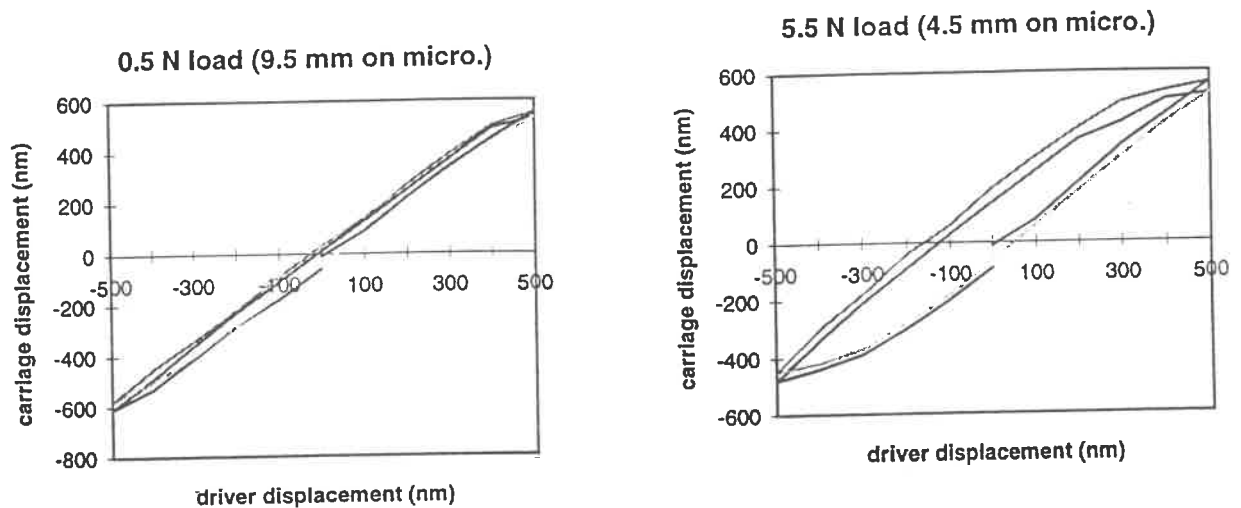


Figure 4: 'Raw' data of displacement resulting when drive is cycled at different bearing loads.

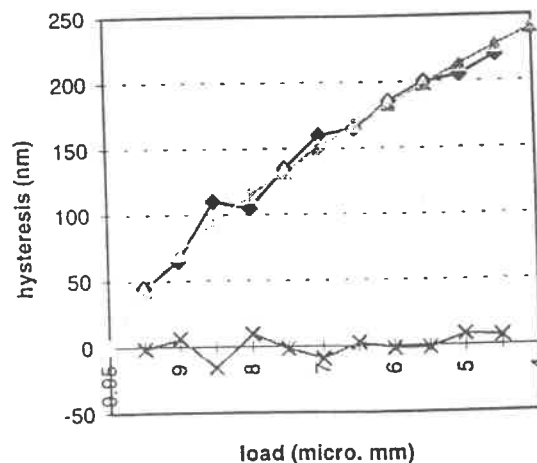


Figure 5: Measured hysteresis and best-fit form of equation (7), with residuals also plotted.

Parameter estimation. While the current experiments can readily show the general functional behaviour, they cannot provide values for the actual interface conditions, since the effects of several variables that cannot be directly determined are combined there. Nevertheless it is instructive to examine the data in the light of plausible assumptions about interface conditions. The combined spring constant of the suspension is about 1000 N m^{-1} . The friction coefficient for PTFE on glass is normally between 0.08 and 0.2. A relatively high value of 0.15 is used here because there is some evidence that the friction coefficient rises slightly at low loads and a very low traverse speed is being used (dynamic coefficient $\approx$ static coefficient). Given such considerations $K \approx 5 \times 10^{-7} \text{ N}^{-2/3} \text{ m}$ is the value obtained experimentally.

The model form is not very sensitive to small changes in the elastic moduli of the materials (cube-root relationships), so if we assume the shear modulus for the interface is that of bulk PTFE, $G = 0.15 \text{ GPa}$, and that the pad has Young's modulus between 0.5 and 1.0 times that of glass, a 20 mm radius pad should show $K \approx 2.5 \times 10^{-5} \text{ N}^{-2/3} \text{ m}$. Comparing this to the experimental value, the shear modulus of the interface needs to be increased by a factor of 50 to match our model. This gives $G = 7.5 \text{ GPa}$, which is around one quarter of that for the substrate metal. This seems not unreasonable: since the PTFE film is very thin the stiffness of the interface is likely to be very different from that of bulk PTFE, but remains, of course, significantly lower than that of the uncoated metal.

Conclusions. Simple modifications to an approximate, analytic method of modelling transverse displacements of Hertzian contacts through microslip appear to be very effective at predicting the functional form of cyclic lost motion in a precision slideway driven by a position asserting system of finite stiffness. Load reduction is shown experimentally to be effective in reducing this loss. The model provides guidance on the relative effect on the lost motion of varying different design parameters.

References.

- [1] Lindsey K., Smith S.T. and Robbie C J, 1988, Sub-nanometre surface texture and profile measurement with Nanosurf 2, *Ann. CIRP*, **37**, 519-22.
- [2] Garratt J. D. and Bottomley S.C., 1990, Technology transfer in the development of a nanotopographic instrument, *Nanotechnology*, **1**, 38-43.
- [3] Mindlin R.D. and Deresiewicz H., 1953, Elastic Spheres in contact Under Varying Oblique Forces, *Trans. ASME; J.Appl.Mech.*, **75**, 327-344.
- [4] Monteiro A.F., Smith S.T. and Chetwynd D.G., 1996, A super-precision linear slideway with angular correction in three axes, *Nanotechnology*, **7**, 27-36.

MICRO-DYNAMICS OF A SLIDE MECHANISM USING ROLLING ELEMENTS

Yuichi Okazaki*, Shun'ichi Sasaki**, Ryoji Murata**

* Mechanical Engineering Laboratory
Namiki 1-2, Tsukuba, Ibaraki, 305 JAPAN

**Science University of Tokyo
Yamazaki 2641, Noda, Chiba, 278 JAPAN

1. Introduction

Precision motion control of scanning stages is one of the key technologies for scanning probe profilers and beam scanning type lithographer for semiconductors aiming to expand the scanning area, as well as for precision NC machine tools. Rolling guides are most commonly used in linear scanning stages taking advantage of their smaller friction compared to sliding guides and their moderate guiding accuracy depending on the geometrical accuracy of the components. On the other hand its nonlinear friction force often disables the accurate motion control, so that the evaluation of the behavior of the mechanism in a small displacement region is important task for finer motion control. It was already found in 1950th that rolling mechanism shows elastic characteristics in a very small displacement region, i.e., that even a small driving force induces displacement, and that the displacement decreases as the force is removed, along a hysteresis curve [1]. Recently a couple of researches have been started focused on the characteristics of rolling mechanisms intending for ultraprecision motion control of mechanisms [2][3]. In most reports the characteristics were measured quasi-statically or based on simple frequency responses. Dynamic measurements and analysis of the nonlinear characteristics of mechanical slide comprising rolling elements have been done, and a coordination of results of static and dynamic tests concerning the behavior.

2. Experimental setup

The linear slide mechanism used is an ordinary designed stage comprising cross-roller guides driven by a direct mounted DC torque motor (Inland: T-1242-D) through a ball-screw of 14 mm diameter. The clearance of the screw is eliminated by preloading. The motor drive is a common velocity control using a tachogenerator. The displacement of the slide was measured by a laser interferometer with 2.5 nm resolution. The relationship between the velocity command, the displacement and the driving torque derived from the drive current of the motor were measured, while sinusoidal signals with variety of amplitude and frequency were applied as the velocity command.

3. Experimental results and discussion

(1) Frequency response

Fig. 1 shows the measured transfer function from velocity command to the displacement. The system behaves as an integrator for large amplitude commands, in the nature of a velocity control system. For a small amplitude command a resonance-like corner frequency was observed, beyond which the gain decay increases almost -60

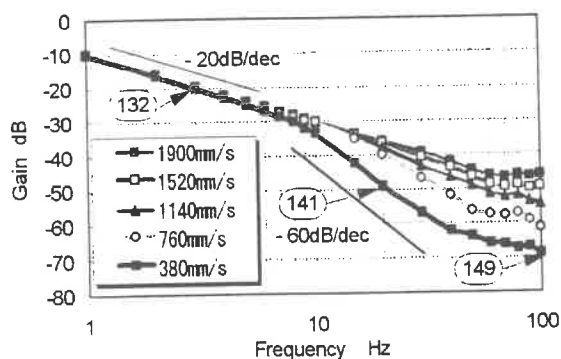


Fig. 1 Frequency response for various input amplitudes

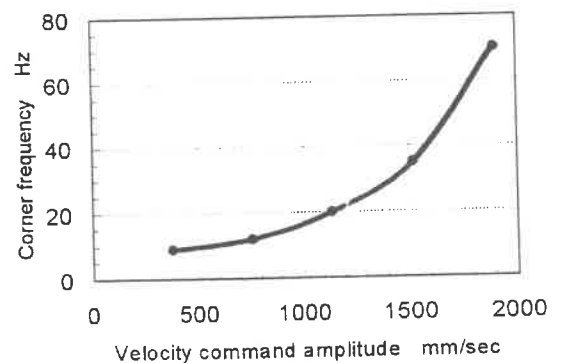


Fig. 2 Resonance-like corner frequency depends on the command amplitude

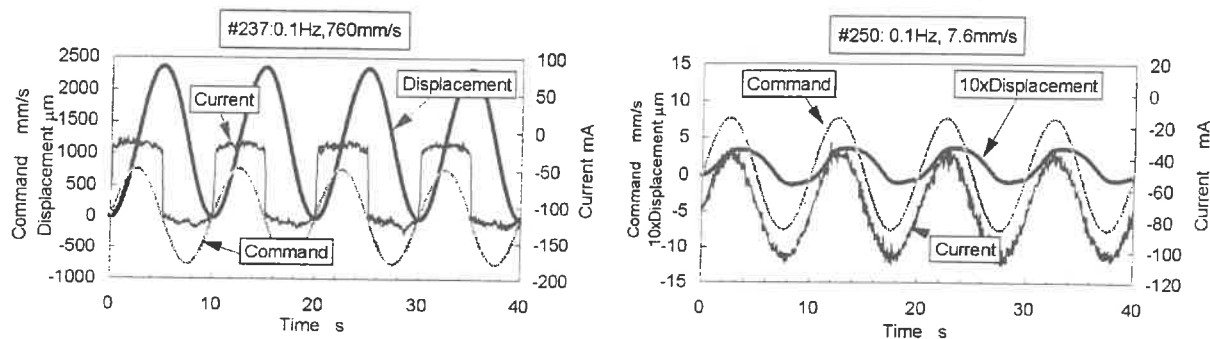


Fig. 3 Command, drive current and displacement.

db/decade. The gain decay is recovered over 60 Hz. The phase lag of the displacement response to the velocity command changes to be -100 degree, -135 degree and -100 degree at 30 Hz (#132), 20 Hz (#141) and 100 Hz (#149) respectively for the command single side amplitude of 380 mm/s. The location of the corner frequency depends on the command amplitude, increasing almost exponentially of command amplitude (Fig. 2). This result corresponds to a preceding report [3].

(2) Drive force and displacement

Fig. 3 shows some time series data of the velocity command, the armature current and the displacement. Under some conditions a the saturation of the drive current occurs, despite that both the command and the displacement change smoothly. Other plots of the current versus the displacement (Fig. 4) explain that as reported in [2] the curve is divided into three regions, i.e. region I: the stiff small displacement region, region II: the transient region, and region III: the large displacement region where the drive current is almost constant. The two boundaries lie around 100 nm and 5 to 10 μm, but the boundary between region II and III is not obvious. These results well correspond to the reported by quasi-static tests [2]. Furthermore, the result of #199 (Fig. 4-c) shows that this relationship is kept for a dynamic drive as well as for a quasi-static drive.

Fig. 5 summarizes the relationship between the phase difference of drive current and the displacement as a function of the command condition. Some data points in preceding figures are marked in the plot. A rough boundary line of the current saturation is also plotted in the figure. For a command of small amplitude and high frequency, the current and the displacement are in-phase, while for a large amplitude and low frequency the current leads the displacement by 90 degrees in phase. Note that a high frequency corresponds to a small displacement amplitude according to the frequency response.

3) Displacement amplitude, resonance-like corner frequency and drive current saturation

Fig. 6 is showing a mapping of the peak-to-peak displacement amplitude for various velocity commands. The command conditions which arises the resonance-like corner frequencies presented in Fig. 2 are also plotted. The corner frequencies are located close to the line of the displacement amplitude = 10 μm. The actual displacement amplitudes at the corner frequencies lie among 9 and 17 μm.

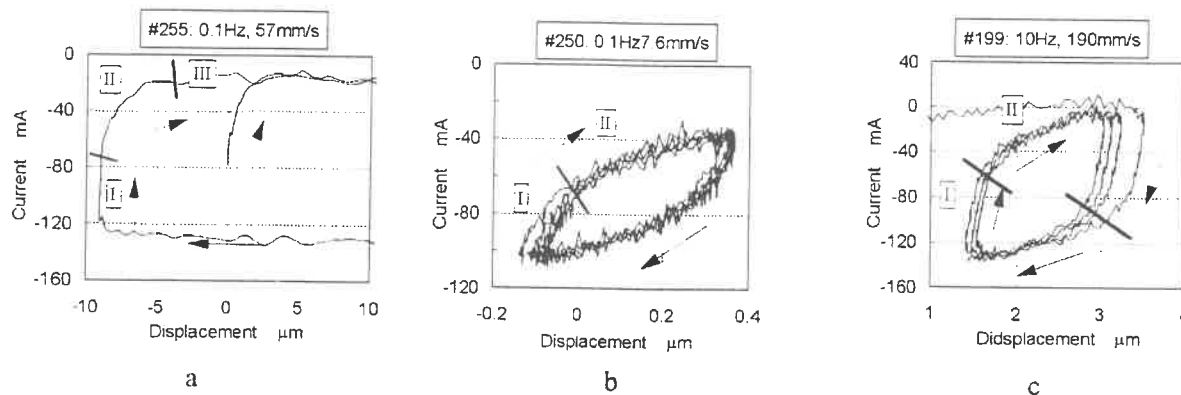


Fig. 4 Some plots of drive current vs. displacement

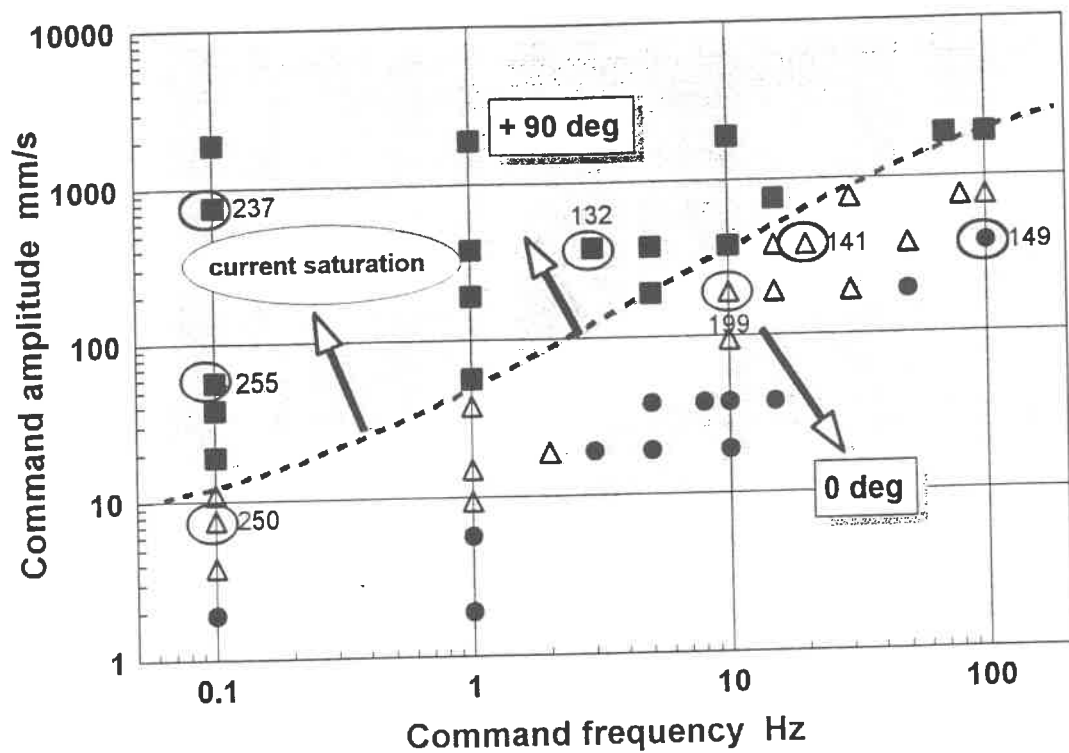


Fig. 5 Phase lead of the drive current to the displacement

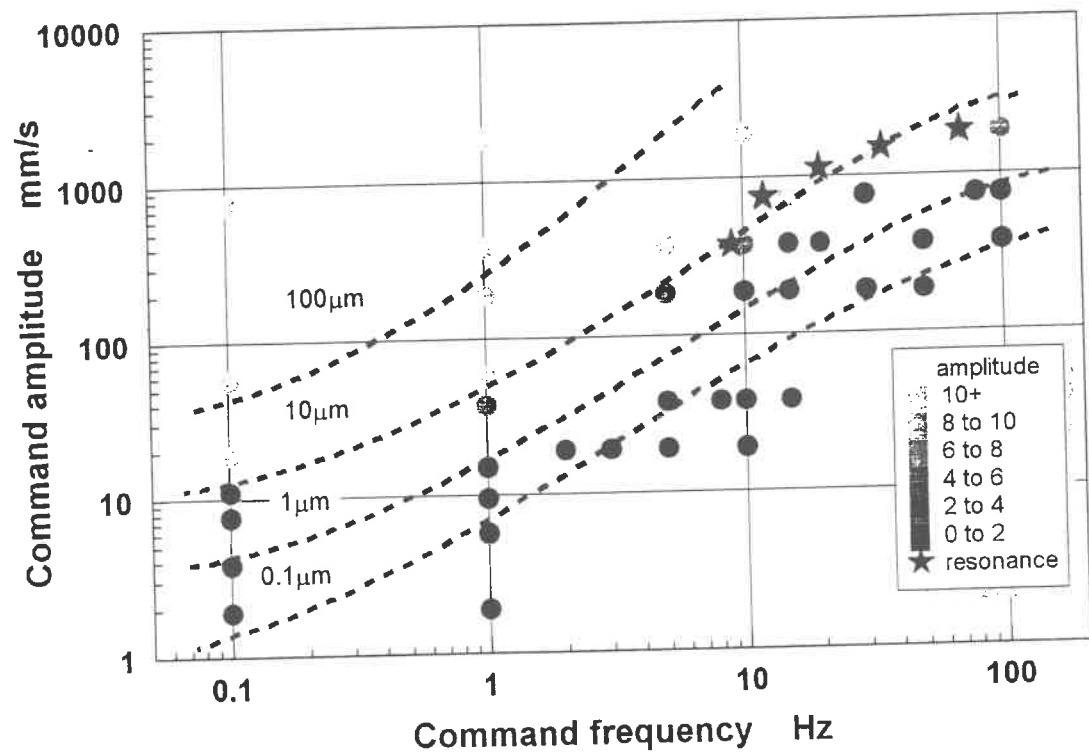


Fig. 6 Relationship between the displacement amplitude and the velocity commands

By comparing Fig. 5 and Fig. 6 the following features have been recognized: The current saturation occurs when the displacement amplitude is about 10 μm . The phase of the current leads that of the displacement by 90 degrees in the region of the displacement greater than 10 μm , and is in-phase when the displacement amplitude is smaller than 0.1 μm . The locations of resonance-like corner frequencies also roughly match the conditions for a constant displacement amplitude of 10 μm . In the region I in Fig. 4 the displacement amplitude is smaller than 0.1 μm . Because in this region the current and the displacement are in-phase, it proves the system is "elastic" there. In the same manner the model for the region III, i.e. that the system is "rolling" with a constant drive force, is also explained based on the phase difference.

By correcting the phase difference between the drive current and the displacement to zero, another features of current-displacement diagrams are revealed (Fig. 7): For almost all tests the correlation of them is virtually linear up to some points, and if the amplitude is large enough, the drive current saturates beyond the points. This saturation should correspond to the saturation in the leading discussion using Fig. 4, but is clearer. These curves composed of a straight line and two saturation lines mean that for small displacement amplitude without saturation, the waveforms of the drive current and the displacement are similar, despite that the plots for region I and II in Fig. 4 are curved. The saturation means that the current waveform is distorted compared to the displacement waveform. The average gradient of the curve represents the ration of the drive current to the displacement. For a large displacement with saturation the ratio is smaller than a small displacement. This typical difference explains the gain decrease over the resonance-like corner frequency in the frequency response.

Note that in this diagram only the waveform is discussed, and that the displacement amplitude between the saturation points does not correspond to the actual displacement, and the gradient of the first straight line does not directly mean the effective stiffness of the elastic region I, because the phase relationship is lost.

Another understanding of Fig. 7 b is that it represents the drive force as a function of the velocity because the phase difference is about 90 degrees under the condition that current saturation occurs.

4. Conclusions

Through dynamic sinusoidal drive tests to observe micro-dynamics of rolling guideway system the following nonlinear characteristics of the mechanism have been revealed:

- 1) As already reported, in the diagram of the drive force vs. the displacement the curve is categorized into three regions by the gradient and whether the displacement is elastic or with rolling. This characteristic is also kept under a dynamic drive up to 100 Hz as well as a quasi-static drive qualitatively and quantitatively.
- 2) The resonance-like corner frequency in frequency response, the phase change and the burst of the displacement amplitude in drive current-displacement transfer characteristics can be associated with in terms of the drive current saturation around 10 μm of the displacement amplitude.
- 3) The models for the three regions are proved by phase difference of the drive current and the displacement.

References

- 1) T. Hisada, Rolling friction and knife-edge resistance. Government Mechanical Laboratory Report 3(1950) (in Japanese)
- 2) S. Futami et. al., Nanometer positioning and its micro-dynamics. Nanotechnology 1 (1990)31
- 3) P. Ro et.al., Nonlinear micro-dynamic behavior of a ball-screw driven slide system. Prec. Eng. 14-4 (1992)229

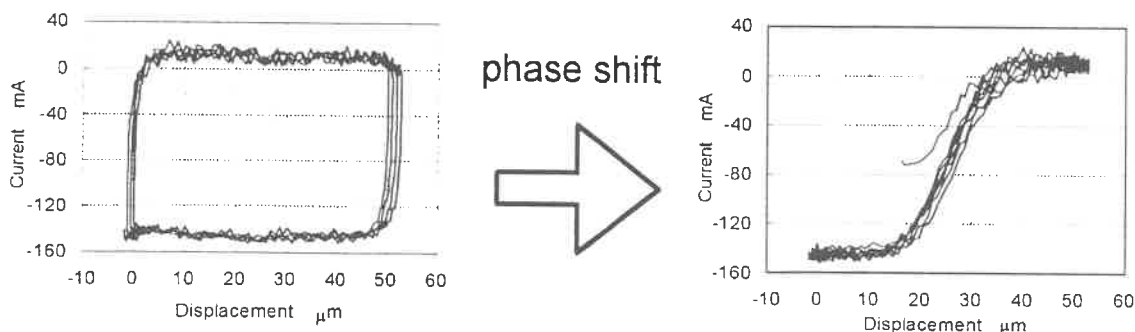


Fig. 7 Phase shift makes the current saturation clear

CHARACTERIZATION OF AN AIR BEARING FOR MIXING STUDIES OF BIOPROCESSES

J.M. Patiño¹, R. Vázquez¹, G. Ascanio¹, E. Galindo² and L. Serrano-Carreón².

¹ Centro de Instrumentos, UNAM, Apdo. Postal 70-186, México, DF 04510

² Instituto de Biotecnología, UNAM, Apdo. Postal 510-3, Cuernavaca, Mor. 62271

Abstract

A new air bearing used for mixing studies bioprocesses has been characterized. This bearing has been improved taking into account the possibility to carry out experiments with higher working volumes. The pneumatic bearing has been designed for using in dynamometer for the accurate measurement of power input and mixing applications in bench stirred reactors.

In this new design many changes have been carried out, such as 18 thrust nozzles distributed in three channels, which provide two resultants: radial and axial, allowing to work with high stability. Results of characterization show a 106.926 kg maximum load when pressurized air at 5×10^5 Pa is introduced. For these conditions, a 2.8×10^{-5} m thrust clearance was found, and the static friction coefficient between air and the dynamic part was 3×10^{-6} .

Introduction

Power input is variable commonly used in chemical and bioprocess engineering. It is defined as the amount of energy necessary in a period of time, in order to generate the movement of the fluid within a container (e.g. biorreactor, mixing tank, chemical reactor, etc.) by means of mechanical or pneumatic agitation.

The costs associated with power input contribute to the overall costs of

industrial plants. Therefore, it is desired that the mixing processes are performed efficiently and with a minimum of energy expenses required to achieve the objective established *a priori*.

The knowledge of the power input, in a specific stirred tank reactor, is necessary for scale-up and scale-down procedures, process characterization, modelling, mixing performance as well as transfer and fluid dynamic studies in a bioprocess. The success of a fermentation process depends strongly on the energy transmitted by the impellers to the liquid in the biorreactor tank, therefore the power input is an important factor in the determination of process costs.

Several methods are available for the measurements of the power input in agitated vessels. On large scale, wattmeters and ammeters are suitable. On pilot scale dynamometers coupled to the shaft are commonly used. On bench scale, torquemeters and air bearings dynamometers have been used. In all scales torsion dynamometers and strain gauges can be used.

Dynamometers with bearing supporting a mixing vessel, have been used for the measurement of the power input in bench scale stirred tank reactors. In this case, the reaction torque due to rotation of the impeller in a specific fluid is determined. Nienow and Miles (1969) have developed an air bearing with a stable and trouble-free operations in which friction losses can be practically neglected.

Pneumatic bearings have several applications at industrial and laboratory levels, for example, as a part of dynamometers to measure the power input to a fluid in mixing processes. Considering its low friction coefficient, pneumatic bearings are suitable devices for this kind of equipments.

Description

A dynamometer has been developed for the accurate measurement of power input and mixing applications in bench stirred tank reactors. The dynamometer consists of a drive system, a transparent vessel and air bearing, which characterization is presented in this paper. The drive system

consists of a direct current motor coupled to a belt-pulleys arrangement and an agitation shaft with one or two impellers, depending on the experiment. The impeller speed is controlled by an electrical potentiometer measured with a magnetic pick-up, a gear and a tachometer. The vessel consists of a cylinder and a water jacket for controlling the process temperature. In one side of the vessel, a rigid arm pushes against a load cell tip transducer placed at right angles to the arm. The torque transmitted by the impellers to the liquid can be measured when the vessel floats on the pneumatic bearing and tends to turn without restriction.

A diagram of torquemeter and ancillary equipment is shown in figure 1.

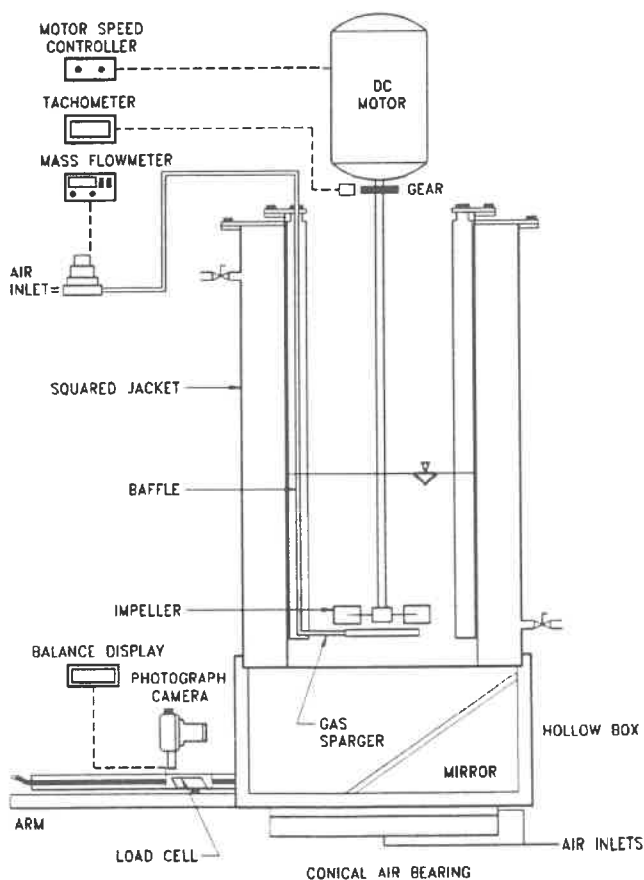


Fig. 1 Diagram of torquemeter with a pneumatic bearing

Design considerations were similar to those originally formulated by Reséndiz *et al* (1991).

The pneumatic bearing has two complementary elements: one static (lower) and the dynamic or floating (upper). The static element has a 40° conical cavity, whose function is to provide axial and radial thrust

component and a central component which provides just radial thrust component to the journal bearing. The pressurized air is introduced into the static element and distributed through the 18 thrust nozzles in three channels, which allows to operate with high stability due to a better air distribution. The dimensions are shown in figure 2.

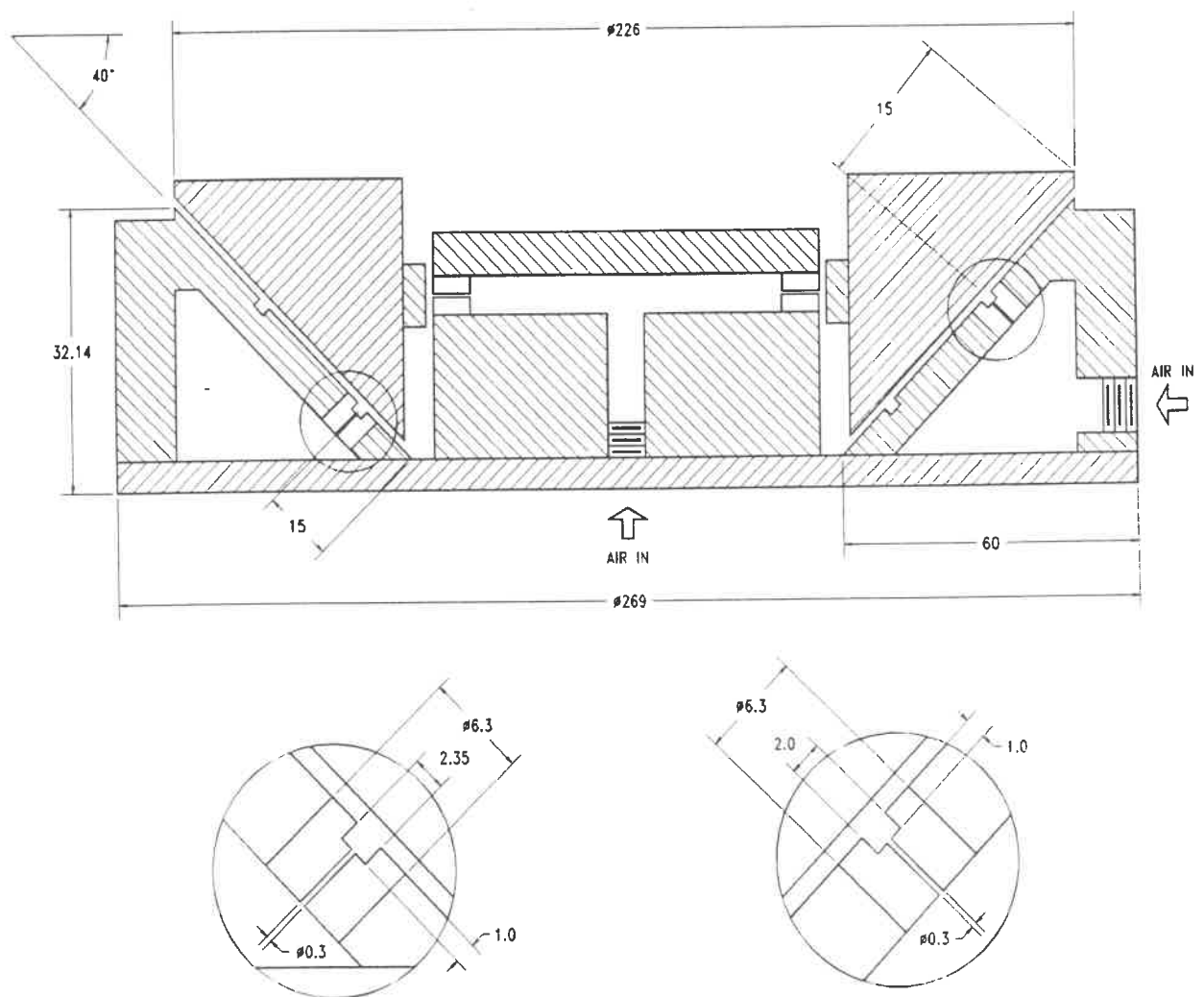


Fig. 2 Dimensions of the conical air bearing (in mm)

Characterization

Using a precision dial indicator, a 2.8×10^{-5} m thrust clearance was found, which is necessary for free rotation to occur between the static and dynamic part, when pressurized air is introduced into the pneumatic bearing.

The minimum torque required to rotate the floating vessel, named static friction torque is 4×10^{-4} Nm. Considering a very low static friction coefficient between the dynamic part and the air (3×10^{-6}) the pneumatic bearing can be considered practically frictionless.

Operating pressures to support masses of 25, 50 and 75 kg were found to be 1.17×10^5 , 2.3×10^5 and 3.5×10^5 Pa, respectively. The maximum load that the pneumatic can support is 106.925 kg when pressurized air at 5×10^5 Pa is introduced into the bearing.

Conclusions

A pneumatic bearing used for bioprocesses applications was developed and characterized. This new air bearing concept was carried out taking into account a previous design of Resendiz *et al* (1991). One of the fundamental differences to those originally formulated is the 40° conicity instead of 30° , as well as the 18 thrust nozzles distributed in three channels, which provide a high stable and trouble-free operation, considering its static friction coefficient, which is of the order of 10^{-6} . The new pneumatic bearing allows to work with higher loads compared to the previous design.

References

- 1 Brown, D., *Chem. Ind. (N.Y.)* 16 (1977) pp. 684-688.
- 2 Nienow, A.W. and Miles, D., *J.Sci.Instrum.* 2 (1969) pp. 994-995
- 3 Reséndiz, R. et-al, *Chem. Eng. Technol.* 14 (1991), pp. 105-108
- 4 Holman, J.P., *Experimental Methods for Engineers*, McGraw-Hill, Tokyo 1971.
- 5 Patiño, J.M., Vazquez, R., Bachelor thesis, National University of Mexico (1995).

Industrial inspection interferometer having a large equivalent wavelength

Peter de Groot and Carl Zaroni

Zygo Corporation, Laurel Brook Road, Middlefield, CT 06455

We have constructed an interferometer suitable for testing flat surfaces that are either too rough or deformed for conventional visible-wavelength interferometry.* The instrument uses a diffractive optical assembly to illuminate the test piece at two different incident angles, using visible wavelength light. The reflected beams recombine to generate an interference pattern with an equivalent wavelength twenty times larger than the actual wavelength of the illumination.^{1,2} The slope tolerance and the acceptable surface roughness are correspondingly increased with respect to conventional interferometry.

The diffractive optical assembly is comprised of two linear diffraction gratings having a one-to-two frequency ratio (Fig.1). The beams make a double pass through the gratings, with a reflection from the object between the passes. The interference fringes have excellent contrast, because both of the interfering beams reflect off the object and the intensity balance is nearly independent of the surface reflectivity. The interference effect is achromatic, and conventional incandescent sources generate clear, colorless fringes.³ The design is also compatible with inexpensive visible-wavelength laser diodes and LED's.

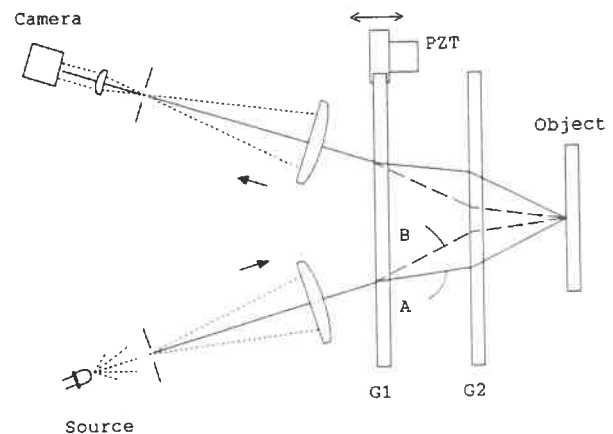


Figure 1: Grating interferometer for flatness testing. The equivalent wavelength is 12 μm .

* US Patent No.5,526,116 (1996). Additional US and Foreign Patents Pending.

A 100-mm aperture prototype instrument uses 100 X 150mm gratings and an equivalent wavelength of 12 μ m. The working distance is 60mm, which is far greater than that of grazing incidence interferometers. The maximum allowable surface roughness, as determined by measuring a sequence of machined metal parts, is 2.5 μ m Ra. This is significantly greater than the 0.3 μ m maximum roughness of the 1.55- μ m laser diode interferometer reported at this meeting three years ago in Seattle.⁴

The maximum allowable peak-valley departure for a purely spherical distortion is 150 μ m over the 100-mm aperture. Phase-shifting techniques provide a single-measurement rms repeatability for smooth surfaces of .25nm and an accuracy of $\pm 0.1\mu$ m. An additional benefit of the interferometer is the ability to separate the front and back surface reflections from transparent, plane-parallel objects.⁵

The 12- μ m equivalent wavelength is a good compromise between measurement range and the increasing demands of precision metrology in manufacturing. The instrument is well suited to the metrology of manufactured items such as engine parts, components for magnetic storage devices, flat-panel displays, molded and textured plastic surfaces, and mechanical pump surfaces and seals (Fig,3). The common-path geometry is insensitive to vibration, and the large working distance will be advantageous for automated optical inspection. Additional information regarding this instrument is available in several publications.^{5,6,7}

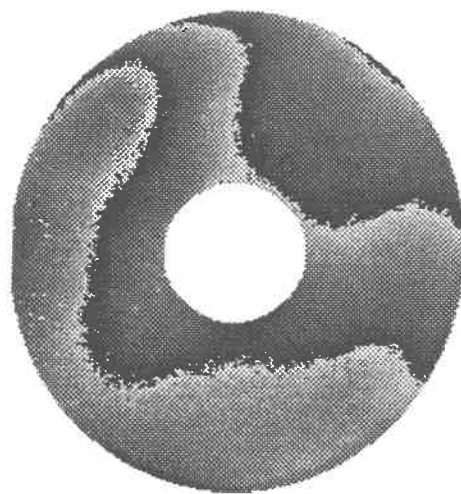


Figure 2: Wrapped phase data for a 95-mm diameter aluminum hard disk substrate. The sensitivity is 6mm of surface departure per fringe. The rms surface roughness is 0.4mm.

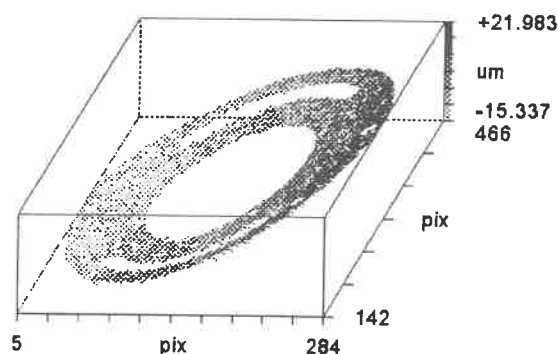


Figure 3: Oblique view of the machined surface of a 70-mm diameter mechanical pump component.

The Authors are pleased to acknowledge the contributions of C. Berg, J. Biegen, T. Connolly, L. Deck and D. Seniff to this work.

References

- ¹ K. Lewotsky, "Grating interferometer measures large departures," *Laser Focus World*, 40-41 (April 1996).
- ² P. Jacquot and P. M. Boone, "Deux methodes holographiques appliquées au controle de planéité," *Proceedings OPTO 88*, 313-319 (ESI Publications, Paris, 1988).
- ³ B. J. Chang, R. Alfemess, and E. N. Leith, *Appl. Opt.* **14**, 1592 (1975).
- ⁴ P. de Groot and L. Deck, "Near-infrared phase-modulation interferometer for surface flatness measurement" *Proc. ASPE*, 33-36 (1993).
- ⁵ P. de Groot, "Grating interferometer for metrology of transparent flats " *Optical Fabrication and Testing Workshop*, OSA 1996 Technical Digest Series, 7, 28-30 (Boston, 1996).
- ⁶ P. de Groot, "Grating interferometer for flatness testing," *Opt. Lett.* **21**(3) 228-230 (1996).
- ⁷ P. de Groot, "Equivalent-wavelength interferometry using diffractive optics " *Proc. SPIE* **2782**-48, *Lasers, Optics, and Vision for Productivity in Manufacturing I* (1996).

APPLICATION OF LIQUID CRYSTAL PANEL AS DIFFRACTION GRATING TO ZONE-PLATE INTERFEROMETER

Kazuhide Kamiya¹, Takashi Nomura¹, Hiroshi Miyashiro¹,
Kazuo Yoshikawa², Hatsuzo Tashiro²

1. Toyama Prefectural University, Kosugi-machi, Toyama, 939-03, Japan

2. Toyama University, Gofuku, Toyama, 930, Japan

Zone-plate interferometers⁽¹⁾ are used to measure an error in the shape of aspherical surfaces. It is not much influenced by vibrations and air turbulence compared with other interferometers, because it has a reference position on a surface under test and is a common path interferometer. A zone plate is a computer generated hologram and generates a diffracted wavefront whose shape is coincident with the design shape of a mirror surface. Therefore the manufacturing the zone plate is needed every time surface under test is changed. The manufacturing process has an exposure by laser, a development and a coating SiO₂⁽²⁾. As the process needs much time, the method of manufacturing zone plate in a short time is needed. Liquid crystal panels have a property of intensity modulators and change displayed patterns in video rate. We propose a zone-plate interferometer using the liquid crystal panel. As the phase of a interference fringe changes proportionally to the phase fraction of the zone plate displayed on liquid crystal panel⁽³⁾, in the proposed interferometer the fringes is analyzed by the phase-shifting method⁽⁴⁾.

Experiments are conducted using a liquid crystal panel from a projection type TV (Epson VPJ-2000) shown in Fig. 1. Specifications of the panel are shown Table 1. Resolution of the liquid crystal panel is high. As an aspect ratio of the panel is 1:1.14, the pattern of zone plate is compensated. A pattern of zone plate generated by computer is shown Fig. 2. The patterns have gray level from 0 to 255 and the

Table 1. Specifications of liquid crystal panel (Epson VPJ-2000 LC Projector).

Type	Twisted nematic, transmission TFT active matrix
Panel size	26.9mm(H) × 20.2mm(V)
Number of pixels	480(H) × 440(V)
Pixel pitch	56 μm(H) × 46 μm(V)

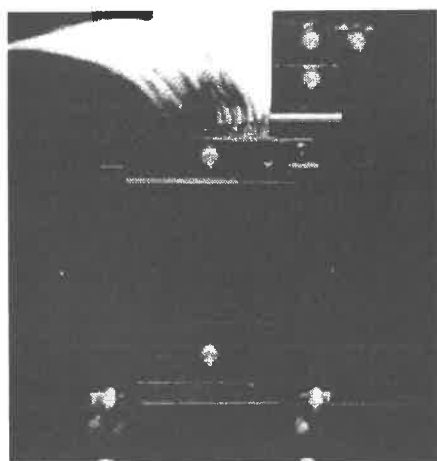


Fig. 1 Photograph of liquid crystal panel (Epson VPJ-2000).

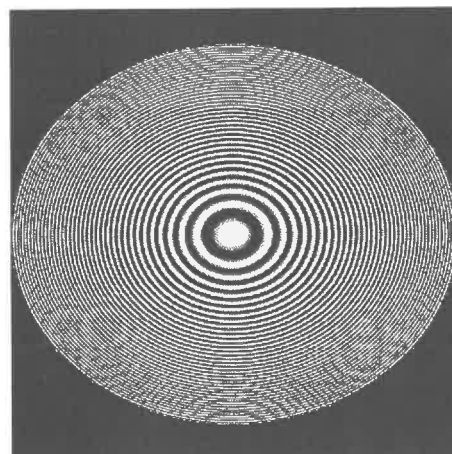


Fig. 2 Pattern of zone plate

aspect ratio of the pattern is compensated.

Flat Mirror is measured by using the liquid crystal panel zone plate. Figure 3 shows schematic diagram of the zone-plate interferometer using liquid crystal panel. Liquid crystal zone plate is illuminated by the parallel incident light. The light through the Liquid crystal zone plate is measurement light. And the diffracted light at liquid

crystal zone plate is reference light. The measurement light reflected by the surface under test is distorted by error in the shape of surface under test and diffract at liquid crystal zone plate. The reference light passes through the liquid crystal zone plate. The both light is converged to a stigmatic point at aperture stop. The extraneous lights are removed almost completely by the aperture stop. The interference fringe is obtained on a charge coupled device camera. The photograph of the experimental setup is shown in Fig. 4. The interference fringe obtained by the proposed interferometer is shown Fig. 5. The

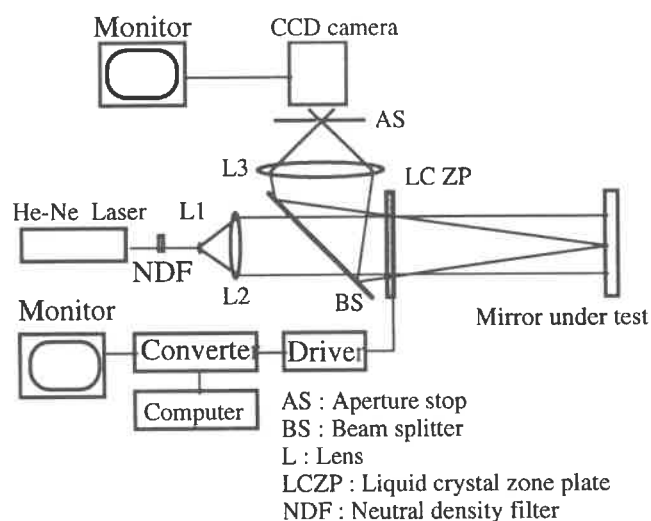


Fig. 3 Schematic diagram of a zone-plate interferometer using liquid crystal panel

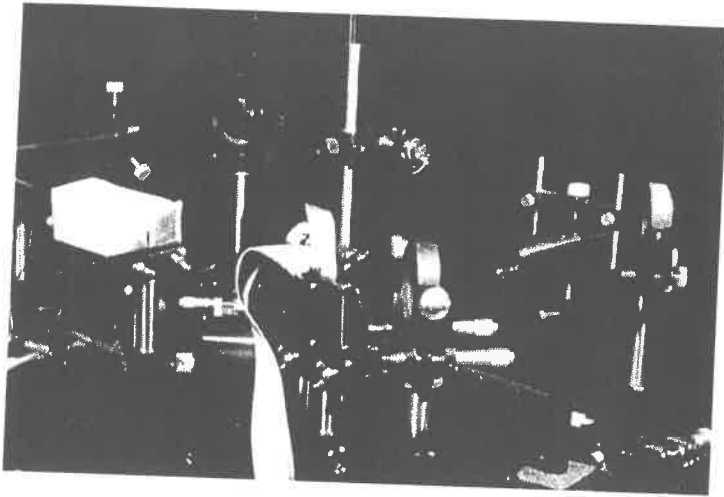


Fig. 4 Photograph of the experimental setup

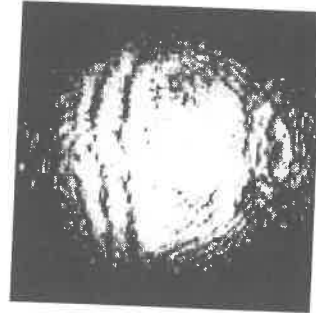


Fig. 5 Interferogram obtained by the zone-plate interferometer using liquid crystal panel.

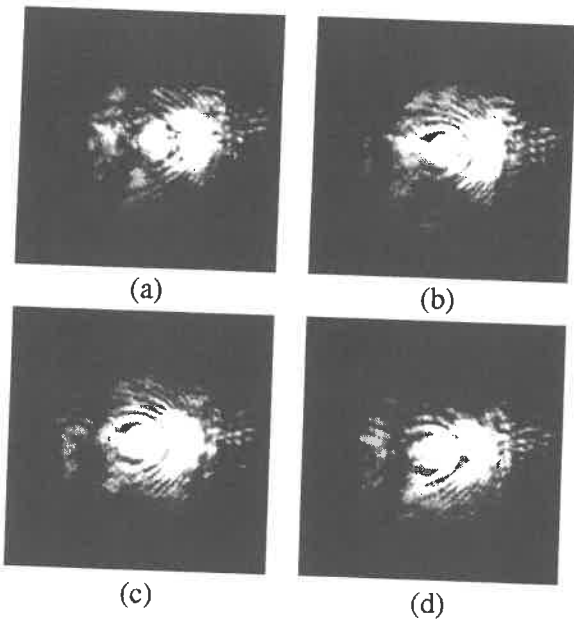


Fig. 6 Interference fringe obtained by the proposed interferometer. (a) Phase modulation of 0, (b) Phase modulation of $\pi/2$, (c) Phase modulation of π , and (d) Phase modulation of $3\pi/2$.

interference fringe is obscure because diffraction efficiency of the liquid crystal zone plate is low. However the interference fringe is stable when the interferometer receive vibrations and air turbulence.

The phase of a interference fringe varies in proportion as the phase fraction of the zone plate displayed on liquid crystal panel. When the initial phase of zone plate is changed to ϕ , the phase of the interference fringe is changed to 2ϕ . Therefore to obtain the fringe with the phase 0, $\pi/2$, π and $3\pi/2$, the initial phases of zone plate are 0, $\pi/4$, $\pi/2$ and $3\pi/4$. Figure 6 shows the fringes obtained by the proposed interferometer. The fringe obtained by the zone-plate interferometer is analyzed by the four step phase-shifting

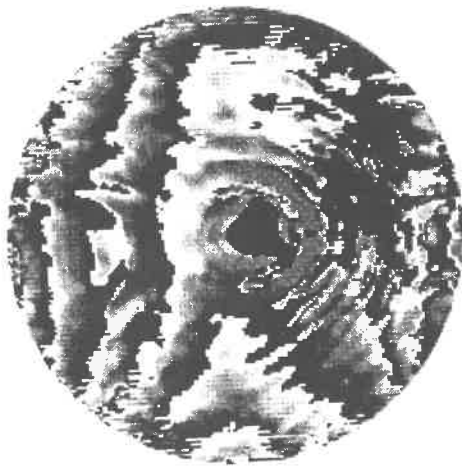


Fig. 7 Phase distribution analyzed by the proposed interferometer

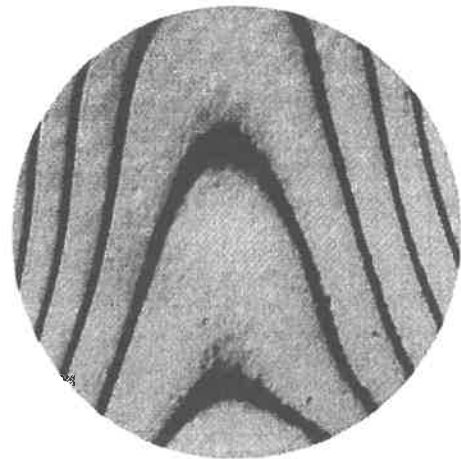


Fig. 8 Interference fringe obtained by Fizeau interferometer

method. The result of analysis from the four fringes obtained by zone-plate interferometer is shown in Fig.7. Figure 8 shows the fringe obtained by Fizeau interferometer to compare the result obtained by proposed method with the result obtained by Fizeau interferometer. The results agree well with the results obtained by Fizeau interferometer.

References

1. R.N.Smartt, "Zone Plate Interferometer," Appl.Opt., 13, 1093(1974).
2. Takashi Nomura, Kazuhide Kamiya, Hiroshi Miyashiro, Kazuo Yoshikawa, Hatsuzo Tashiro, Masane Suzuki, Shigeo Ozono, Fumio Kobayashi, and Masao Usuki, "An instrument for manufacturing zone plates by using a lathe," Prec.Eng., 16, 290(1994).
3. Takashi Nomura, Kazuhide Kamiya, Hiroshi Miyashiro, Kazuo Yoshikawa, Hatsuzo Tashiro, Shigeo Ozono, Masane Suzuki, "Application of a fringe scanning method to zone-plate interferometry," Proc.SPIE, 1720, 317(1992).
4. J. H. Bruning, D.R. Herriott, "Digital Wavefront Measuring Interferometer for Testing Optical Surfaces and Lenses," Appl. Opt. Vol.13, No.11, 2693-2703(1974).

KEYWORDS : zone-plate interferometer, shape measurement, liquid crystal panel, fringe scanning method

Wavelength-Shift Interferometry: Using a Dither to Improve Accuracy

Jack A. Stone, Alois Stejskal*, and Lowell Howard

National Institute of Standards and Technology, Gaithersburg, MD, 20899

*Permanent Address: Institute Of Scientific Instruments, AVCR, Brno, Czech Republic

Introduction: A dither (modulation) in path length can dramatically improve the accuracy of wavelength-shift methods used for absolute distance interferometry. Here we report how a dither improves the accuracy of absolute distance measurement by two orders of magnitude, reducing uncertainty from ~ 1 mm to ~ 10 μ m (1σ). Our results indicate that, when the path length is dithered, residual errors in fringe interpolation can be reduced to less than 0.6 nm after averaging.

"Absolute interferometry" refers to the use of interferometric techniques for determining the position of an object directly, without the necessity of measuring continuous displacements between points. In contrast to a standard interferometer, which measures a *displacement* by counting fringes as the length of the measurement arm is changed, an absolute interferometer directly measures the difference in *length* of its reference and measurement arms, with the length of the arms held constant. We have been investigating a wavelength shifting technique^{1,2,3,4} for absolute interferometry. Using this technique, the position of a cube corner reflector in a single-pass interferometer is determined by observing the phase change at the interferometer output as wavelength of the laser is changed smoothly and continuously between two values. When the wavelength is changed, the relative phase change in light traversing the reference and measurement arms of the interferometer (or the number of fringes passing the detector) is proportional to the product of (1) the difference in length of the reference and measurement arms and (2) the amount by which the frequency of the laser source is shifted. If the frequency shift (or equivalently, the corresponding wavelength shift) is somehow determined, then the position of the cube corner can be inferred from the measured number of fringes passing the detector. In our experiment, the wavelength shift is not measured directly. Rather, the unknown length of the measurement arm (more precisely, the difference between this length and the fixed-length reference arm) is compared to the previously determined length of a "standard" interferometer, similar in construction to the interferometer of unknown length.⁴ The ratio of the length of the unknown interferometer to that of the standard interferometer is simply equal to the ratio of the number of fringes passing in the two interferometers. The actual apparatus is shown schematically in Figure 1.

When using this technique for wavelength sweeping, it is necessary that the diode wavelength is tuned smoothly and continuously, without mode hops. Typically this is possible only over a

narrow tuning range. We obtained first results using a commercial laser for which the range of continuous tuning was only $\Delta\lambda/\lambda \approx 6 \times 10^{-5}$. The number of fringes that are counted is then a factor $\sim 6 \times 10^5$ smaller than would be counted when measuring the distance by standard (displacement measurement) methods. Errors in fringe interpolation are magnified by the reciprocal of this same factor, or 1.7×10^4 , when computing the actual length.

Thus, in order to achieve an accuracy of a few micrometers when using this narrow sweep range, it is necessary to divide a fringe with sub-nanometer accuracy. Achieving sub-nanometer accuracy is difficult primarily because of errors that are periodic or quasi-periodic with wavelength, such as quantization errors, electronic interpolation errors, some forms of coherent scattering errors, and polarization mixing in heterodyne interferometers (not a problem with our geometry). Any of these error sources can be fatal to the wavelength sweeping technique unless appropriate measures are taken to reduce their influence. The effect of periodic errors can be greatly reduced by uniform averaging over all possible values of the phase at the start of the wavelength sweep. We dither this phase by modulating the path length; the cube corner in the reference arm of each interferometer is mounted on a simple loudspeaker and modulated with a triangular waveform with an amplitude of 5 to 25 μm .

Description of Experiment: We used a heterodyne method for fringe counting, with an acousto-optic modulator (AOM) that generates an 80 MHz frequency shift in one arm of each interferometer. The AOM frequency and ~ 80 MHz signals from the two interferometers are mixed down to ~ 2 MHz before processing. In the experiments described here, we processed the down-converted signals using commercial fringe-counting electronics modified so that the direction of counting can be switched synchronously with changes in the direction of the frequency sweep of our laser. As a consequence of the switching, counts that are accumulated when the laser is sweeping up in frequency run in the same direction as counts accumulated when sweeping down, and thus the total accumulated count reflects the average rate of counting for the two directions of frequency sweep.

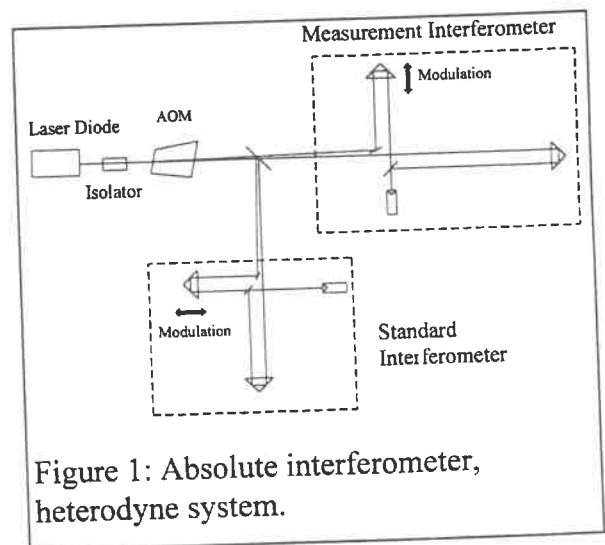


Figure 1: Absolute interferometer, heterodyne system.

In order to assess nonlinear errors in our absolute interferometer, we compared displacements, as determined from position measurements made with our system, to displacements measured using a commercial interferometer. Deviations from a linear relationship between the measurements of the two systems are shown in Figure 2. The data can be interpreted as follows: a linear least-squares fit is used to determine the optical length of our "standard" interferometer at ambient conditions, and the deviations shown (arising from nonlinearities or noise) are errors of our

absolute system excluding overall errors in determining the length of the “standard” interferometer. The uncertainty in the displacement measurement with the commercial interferometer is negligible relative to the errors shown here. This data was taken using a commercial 670 nm tunable laser source with a frequency sweep of 28 GHz ($\Delta\lambda/\lambda \approx 6 \times 10^{-5}$). We were able to achieve 9 μm standard uncertainty (1σ) over a range 0-1.5 m.

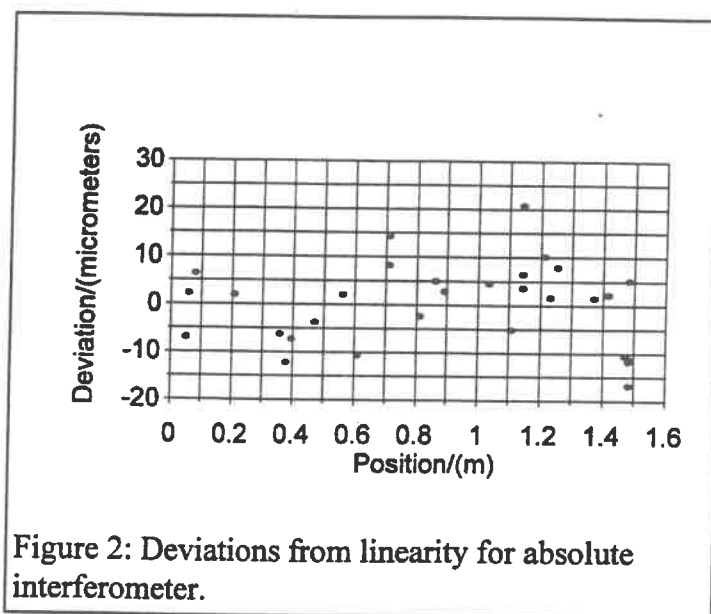


Figure 2: Deviations from linearity for absolute interferometer.

Discussion: In our system, the primary source of error is quantization error. We will discuss mainly this error source, although similar considerations apply to other quasi-periodic errors as well. Our electronics has a resolution of only $\lambda/4$. Without further measures, this will produce errors in length measurement on the order of $\lambda_{\text{syn}}/4$ in the unaveraged result, where λ_{syn} is the “synthetic wavelength”; for a wavelength sweep from initial wavelength λ_1 to final wavelength λ_2 , the synthetic wavelength is $\lambda_{\text{syn}} = \lambda_1 \lambda_2 / \Delta\lambda$. Here λ_{syn} is a factor 1.7×10^4 bigger than λ itself. More precisely, the expected 1σ uncertainty of an unaveraged length reading for either the standard or measurement interferometer is $(\lambda_{\text{syn}}/4)/\sqrt{6}$. This result is a consequence of the fact that the 1σ uncertainty arising from a quantization unit r is $r/\sqrt{12}$, (uniform rectangular uncertainty distribution)⁵, and two independent phase measurements (at the two ends of the wavelength sweep) are required for 1 determination of the phase change. When both the standard and measurement interferometers are comparable in length, the expected uncertainty in the unknown length would be a factor $\sqrt{2}$ larger than this because it is derived from the ratio of two measured phase changes. An uncertainty of about 1.7 mm would be expected for the data shown here if averaging was not used. The observed uncertainty of 9 μm is almost 200x smaller, implying that the effect of quantization error has been correspondingly reduced. To achieve comparable results without averaging it would be necessary to reduce interpolation errors to less than 0.6 nm.

The reduction in quantization error is a consequence of modulation and averaging. We have modeled the measurement in a single interferometer (either the “measurement” or “standard” interferometer) and find the following results. The quantization error is very nearly periodic and averages to approximately zero over 1 unit of resolution ($\lambda/4$). (The error is more precisely periodic over the much longer distance $\lambda_{\text{syn}}/4$.) Uniform time averaging of results will give zero error if the length is modulated using a triangular waveform with peak-to-peak amplitude $n \cdot (\lambda/4)$, where n is an integer. If n is not an integer, uniform time averaging will not produce a uniform average over all quantization errors and some residual error will remain. For a modulation with peak-to-peak amplitude $(n+\epsilon) \cdot (\lambda/4)$, the error has local maxima when $\epsilon \approx 1/2$. The magnitude of these maxima is inversely proportional to the modulation amplitude. Because

the amplitude cannot be set exactly, it is advisable to use a sufficiently large value for n that the error will remain small even if $\epsilon=1/2$. For $n>30$ it is possible to guarantee that the quantization error will be reduced by at least a factor of 100. However, a potential danger of using large-amplitude modulation is described below.

In fact, uniform time averaging is never achieved. Phases are measured only at the extremes of the frequency sweep. Thus the phase is sampled with a sampling rate equal to twice the frequency of the laser frequency modulation, and sampling may give rise to additional problems analogous to aliasing. Suppose, for example, that the length is modulated at a frequency that is exactly $1/10$ the frequency of laser modulation, so that the phase is sampled at 10 equally spaced points in time as the length is modulated $1/2$ cycle. The same 10 points will be sampled during the second half-cycle of length modulation and on all subsequent cycles. If the peak-to-peak amplitude of modulation is 5 resolution units ($5*\lambda/4$), then quantization errors will repeat on every other sample. All the phase-difference measurements, derived from pairs of phase measurements, will sample identical quantization errors, and averaging results with this modulation will not diminish the effects of quantization error over what occurs without modulation. This example is an extreme case, unrealistic because it assumes a precise amplitude that gives particularly bad results and because it requires that the ratio of the two frequencies be an exact integer, which would not occur unless the two modulation signals are somehow locked to each other. When the frequencies are not in an exact ratio of small integers, sampling effects will be minimal after long averaging times, but they may still degrade accuracy at shorter times.

If the peak-to-peak amplitude of modulation were reduced to $1*(\lambda/4)$ unit in the example above, then five distinct quantization errors are sampled when phase differences are measured, and the effect of modulation is essentially equivalent to increasing the resolution of the system by a factor of 5; errors will be 5x smaller than errors obtained without modulation, and the spatial period of the errors will also be reduced by a factor of 5.

Potential sampling problems are most easily avoided by setting the period of length modulation much longer than the period of laser frequency modulation, so that many samples occur during one modulation period. Alternately, if the periods of length and frequency modulation are not in a ratio of two small whole numbers, then the positions sampled will repeat only after a large number of cycles and once again sampling problems will be diminished.

The analysis above avoids the issue of noise. A small amount of noise is beneficial in reducing quantization errors in the same manner as the dither; Gaussian noise is in some respects superior to linear dither and has been used to improve electrical measurements⁶. Figure 3 shows results of a computer simulation of quantization errors under various circumstances, including the presence of noise. On both axes the units are $\lambda_{syn}/4$, the effective resolution of the device when the basic interferometer resolution is $\lambda/4$. Errors are shown for measurements of lengths over a range of about one wavelength, small relative to the synthetic wavelength. The lengths being measured in this example are nearly equal to $0.2*\lambda_{syn}/4$, and essentially the same pattern will be repeated at $(n+0.2)*\lambda_{syn}/4$ for any integer n . The basic quantization error in length measurement is shown by curve (a), a rectangular waveform varying between -0.2 and +0.8 units. Curve (b) in the figure shows the effects of adding random amplitude fluctuations--a Gaussian distribution

with $1\sigma = 20\text{ nm}$ --to each phase measurement, and averaging over 6000 samples (60 s averages if the laser is modulated at 50 Hz). Here the noise reduces quantization errors by about a factor of 2. With both Gaussian noise and a modulation amplitude of $5.1\text{ }\mu\text{m}$ ($30.5\lambda/4$), errors are reduced substantially [curve (c), with the errors multiplied by 10 for clarity.] When noise is eliminated and modulation is present, the errors decrease by more than a factor of 2 relative to curve (c), indicating that in this regime the noise is a hindrance rather than helpful in reducing errors.

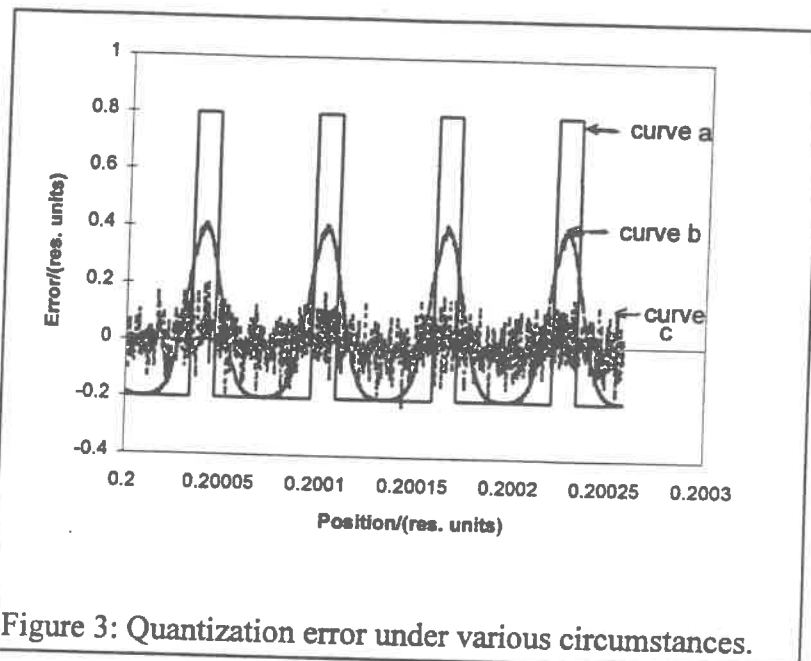


Figure 3: Quantization error under various circumstances.

Concluding Remarks: The computer simulations excluding noise suggest that the dither should reduce fringe interpolation errors somewhat more than observed. Remaining errors that dominate our result when the dither is present are presumably not periodic with fringe spacing. The sources of these errors have not been fully identified, although it is plausible that random vibrational motions of the interferometer reflectors are an important limitation on accuracy when results are averaged for less than 1 minute. In addition, signal processing delays were an important source of errors for short averaging times. We have now built new electronics to reduce these delays, and we have obtained a diode laser with much greater tuning range.¹ We still use modulation to improve our results, but with the larger tuning range the improvement in accuracy when using modulation is not quite so dramatic. The tuning range is now increased by a factor of 30, the effect of periodic errors is 30x smaller, and other sources of error now dominate our results, which have been improved by about a factor of 6 relative to the results presented here.

It may seem disturbing that, although the dither provides averaged phase measurements with sub-nanometer fringe interpolation, the dither itself is thousands of nanometers in amplitude and may introduce drifts in the average length that are larger than the errors that are eliminated. This apparent problem is not so bad as it might appear at first glance. Although it may be true that the dither introduces small errors, these errors are not periodic with the period of wavelength sweeping and consequently can be reduced by averaging.

Our dithering technique of eliminating periodic errors may also be applicable to certain other measurements, such as measurement of linewidth. Using our absolute interferometer to measure a length of 1 meter with $10\text{ }\mu\text{m}$ accuracy is entirely equivalent, from the standpoint of the actual

phase measurements required, to a linewidth measurement that must determine a displacement of 60 μm with 0.6 nm accuracy. The change in optical phase is identical for these two measurements, although in one case the phase change results from tuning laser wavelength by $\Delta\lambda/\lambda=6\times 10^{-5}$ whereas in the other phase change results from motion of a cube corner reflector.

- ¹ J. A. Stone, L. Howard, A. Stejskal, M. Stephens, C. Oates, and L. Hollberg, "Novel Methods for Length Measurement Employing Diode Lasers", Proc. 1996 NCSL Symposium, Monterey, 1996, p 667.
- ² Hisao Kikuta, Koichi Iwata, and Ryo Nagata, "Distance measurement by wavelength shift of laser diode light", Appl. Opt. 25 (17), 1986, pp.976-980.
- ³ Hisao Kikuta and Ryo Nagata, "Absolute distance measurement by wavelength shift interferometry with a laser diode: some systematic error sources", Appl. Opt. 26 (9), 1987, pp.1654-1660.
- ⁴ J Thiel, T. Pfeifer, and C. Haas, "Absolute interferometric distance measurement with tunable laser diodes", VDI Berichte 1118, 3rd Int. IMEKO Symp. on Laser Metrology, Heidelberg, 1994, p.79.
- ⁵ Guide to the expression of uncertainty in measurement, International Organization for Standardization, Case postale 56, CH-1211 Geneve 20, Switzerland, 1995, p.13.
- ⁶ M.F. Wagdy, "Effects of various dither forms on quantization errors of ideal A/D converters", IEEE Trans. on Instr. and Meas., 38, (4), 1989, p.850

ANGLE MEASUREMENT USING X-RAY MOIRÉ FRINGES

N. O. Krylova, D. G. Chetwynd and D. K. Bowen

Centre for Nanotechnology and Microengineering
Department of Engineering,
University of Warwick, Coventry CV4 7AL UK

Introduction and Principles. It is now established that x-ray interferometer (XRI) displacement fringes can provide practical, highly traceable displacements calibration in the nanometre to micrometre region [1-3], with resolution of a few picometres. The method is one used for basic calibration, not for direct application as an industrial gauge, since it is relatively slow and somewhat restricted by the need for x-ray safety enclosures. Nonetheless it is very attractive since the naturally occurring and extremely well characterized lattice parameters of pure single-crystal silicon, known and traceable to $3 \cdot 10^8$ are used directly as secondary standards of length with a fundamental unit of a few tenths of a nanometre. For a review of X-ray interferometry and its applications see [4].

The monolithic XRI is machined from a high-purity single crystal of silicon and then etched to remove surface damaged layers. Three thin parallel 'blades' are left standing proud of the base, oriented so that suitable lattice planes of the crystal are normal to the surfaces and in identical orientation in the three blades. A collimated x-ray beam illuminates the first (splitter) blade at the Bragg angle and diffraction causes two transmitted beams to emerge from the blade at plus and minus this angle. These beams meet the second (mirror) blade where they are each split again; the diverging output beams are masked off and the converging ones meet and interfere inside the third (analyser) blade. Since this interference pattern and the lattice have similar wavelengths, they interfere further such that the phases of the two diffracted beams emerging from the third blade contain information about the lattice positions in all three blades. Displacing the blade in its plane (via a flexure mechanism machined into the monolith) modulates this output intensity to give displacement fringes; one sinusoidal cycle corresponds to a movement of one lattice spacing (around 0.2 or 0.3 nm for the most commonly used lattice orientations).

Very precise calibration of small (sub-microradian) angular displacements is needed in many fields such as precision optics, astronomy and space communications. Interest in XRI for this arises from its proven sensitivity and traceability for linear applications. An angle measuring XRI has been produced by Windisch and Becker at PTB [5] which has sub-nanoradian sensitivity, but its traceability (absolute accuracy) depends directly on geometrical lengths in the monolith (not so easily determined) and its range is restricted by potential errors inherent to its x-ray optical configuration. Another approach is to exploit the moiré fringes generated if one XRI blade is twisted slightly with respect to the others - just as happens with optical diffraction gratings. These fringes have a pitch (wavelength) given by $\lambda = d/\theta$ where d is the lattice parameter. Sensitivity to angle is thus very high (0.1 μ rad twist gives fringes of the order of 1 mm) and the inverse relationship means that errors in fringe measurement are less critical at smaller angles, where fewer alternative methods are available. Traditionally, accidental generation of such fringes from parasitic rotations in a displacement interferometer has been viewed as a problem causing loss of contrast in the displacement fringes, although contrast variation has been used to monitor the angular parasitics of a mechanism [6]. Modern x-ray imaging allows relatively rapid and

cost-effective capture of moiré images and so of directly analysing the pattern for micro-angle measurement. We report on an experimental system to demonstrate the feasibility of this idea.

Measurement System. The system has four main parts: x-ray source, special monolith and associated drive, CCD camera system, and image processing and signal analysis computing. Figure 1 illustrates their arrangement in a set-up for calibrating an autocollimator. The source is a Mo K_{α} sealed tube with a 0.4 mm square focal spot in projection, operating typically at 45 kV and 45 mA. This is coupled to a brass collimator giving an output beam approximately 10×0.4 mm, expanded to 2 mm wide by asymmetric reflection from a monochromator crystal. The tube is 600 mm away from the monolith. A set of movable, and gauged, slits for image calibration is placed between the monolith and imager, which are 100 mm apart. The imager is a conventional low-light CCD camera (753 x 576 square pixels in standard 4:3 video format) fitted with a fibre optic (magnifying) faceplate that is coated with a gadolinium oxysulphide scintillation layer and then a metallic aluminium layer to exclude light. It has a window of 11 mm diagonal and a resolution of about $18 \mu\text{m}$. Thresholded binary images are captured and averaged by a Matrox MVP500 image analysis board in a 486-based PC. The acquisition rate is only ~ 5% less than full frame rate because of data-transfer overheads: good quality images generally involve a 1000 frame average which takes around one minute to obtain, and weaker images rarely need more than 5000 frames.

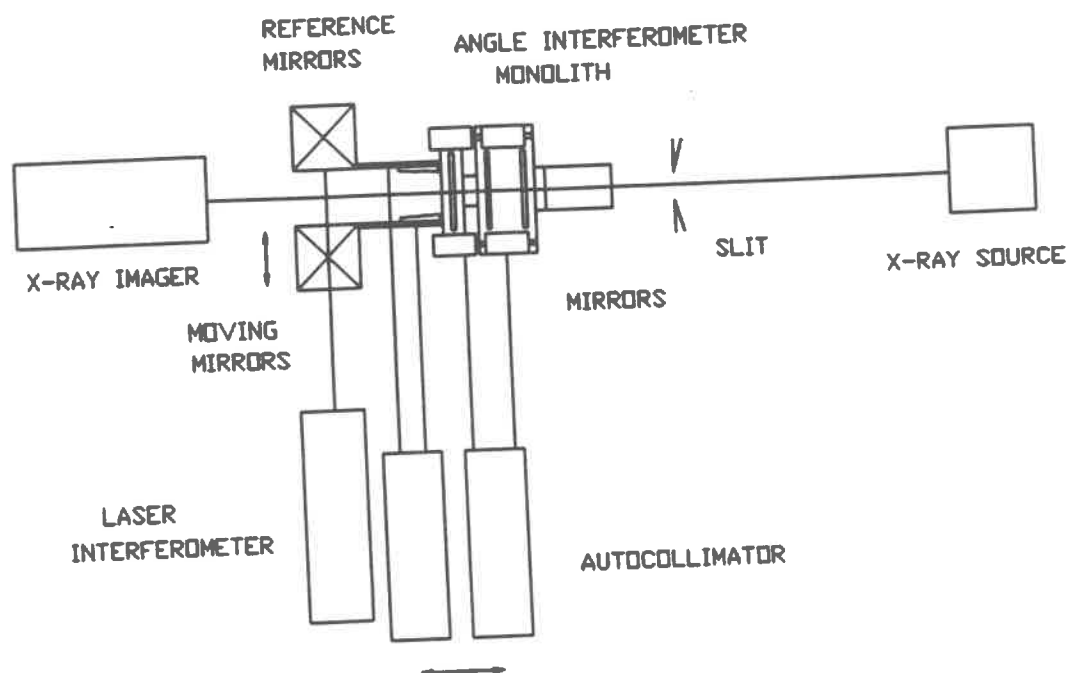


Figure 1: Schematic of an angular X-ray interferometer for testing autocollimators

The new monolith design, figure 2, is the heart of the system. The blades stand vertically on its top surface, each $35 \times 30 \times 1.6$ mm and spaced 15 mm apart. These dimensions provide a 'thick-blade case' transmission interferometer. The crystal orientation gives (220) reflection planes parallel to both long edges of the blade. This is one of the strongest reflections, has the symmetry needed to align to both blade edges and has the added advantage of allowing alternative ranges of angle to be measured by using the (weaker) (440) or even (880) reflections. A problem could arise if the method showed significant sensitivity to displacement fringes superimposed on the moiré fringes since in practice it is

difficult to ensure that the centre of rotation is ideally sited. Our design provides for testing such sensitivity by deliberately placing the centre of rotation about 35 mm below the blade centre to give a considerable Abbe offset and allowing moirés to be taken from the vertical or horizontal crystal planes. The vertical lattice planes experience much larger normal translations than the horizontal ones for a given rotation (the ratio is closely half the radian angle of rotation), so measurements in both planes can detect effects due to displacement fringes. Rotation of the final blade is by applying small forces to the horizontal arm above a single notch hinge. This has nominally 5 mm radius, 10 mm depth and 3.5 mm central thickness, to give a design stiffness of around $1000 \text{ N.m rad}^{-1}$, dependent upon the Young's modulus in the particular crystal orientations used. The monolith incorporates provision for an asymmetric beam expander and monochromator, also adjustable on a notch hinge. Other cuts are for stress diversion, including regions where mirrors may be mounted to the block, to minimize lattice distortion in the blades.

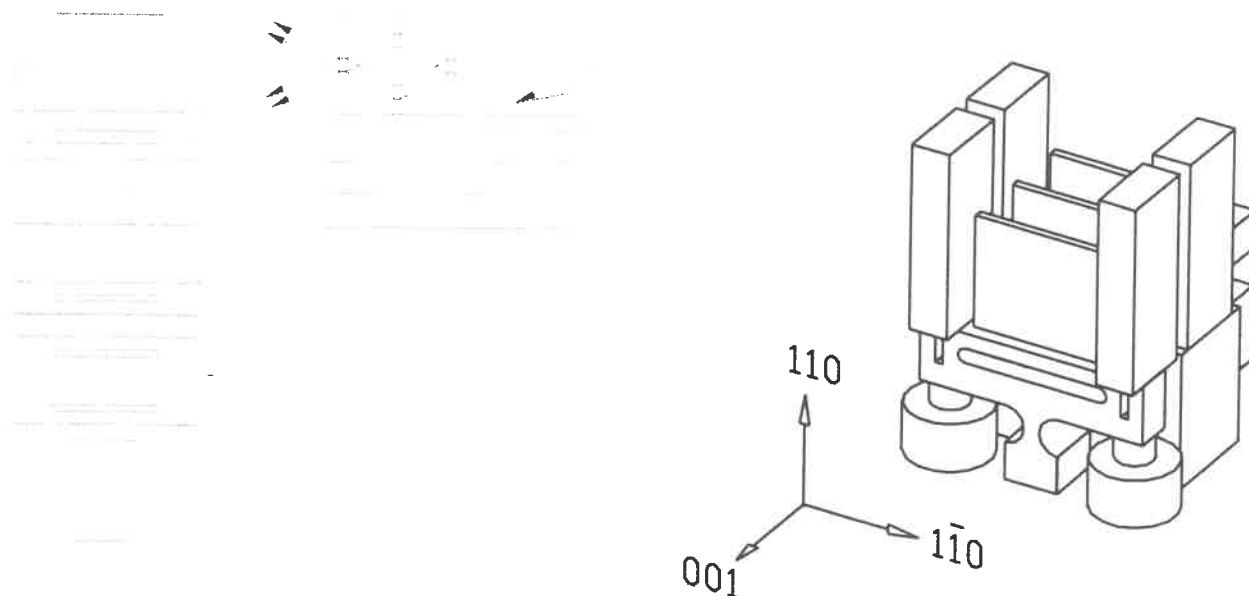


Figure 2: The new monolith, showing typical X-ray paths on the elevation and mirrors and drives on the isometric view.

Blade rotation is driven by small magnet-coil systems in which saturated permanent magnets are glued to the torque-arm of the monolith, surrounded by solenoid coils fixed to the same base-plate on which the monolith rests kinematically. Such actuators are very stable, capable of very high linear range to resolution and have very high compliance so that almost no vibration is transmitted through them [3]. General vibration isolation is provided by standing the whole of this system on a sandwich of thick plywood sheets separated by partly-inflated small pneumatic tyres, which rests on a heavy and actively air-suspended optical bench within a lead/plywood enclosure that provides both radiation and acoustic isolation.

The accurate determination of the spacing of the moiré fringes is of critical concern if the system is to be used for angle calibration. Provided that good fringe contrast is obtained there are several relevant image analysis techniques well known from optical interferometry that are useful under

different conditions. As this paper is concerned with the instrument design, no further discussion will be provided here. However, for any form of analysis the real image size must be accurately determined: the magnification from the final blade, where the fringes are true size, to the camera may not be exactly unity and there may be uncertainty on pixel size. Our approach is to place x-ray opaque artefacts very close to the blade and then to calibrate from their shadow images. For example, a steel sphere placed by the blade can check image distortion. Generally a slit with well-defined edges is kept close to the blade, see figure 1. One edge is fixed to one side of the field and the other is on a precision transverse slide, position monitored by laser interferometer (or inductive gauges in some cases). Fringe images taken with the moving edge in two positions of closely known separation give immediate data on the effective magnification. Linearity checks across the field may also be made. Often the slit spacing (which is always slightly less well defined than any one displacement) is sufficient alone for confirming that calibration is being maintained and the interferometer need not be used continuously.

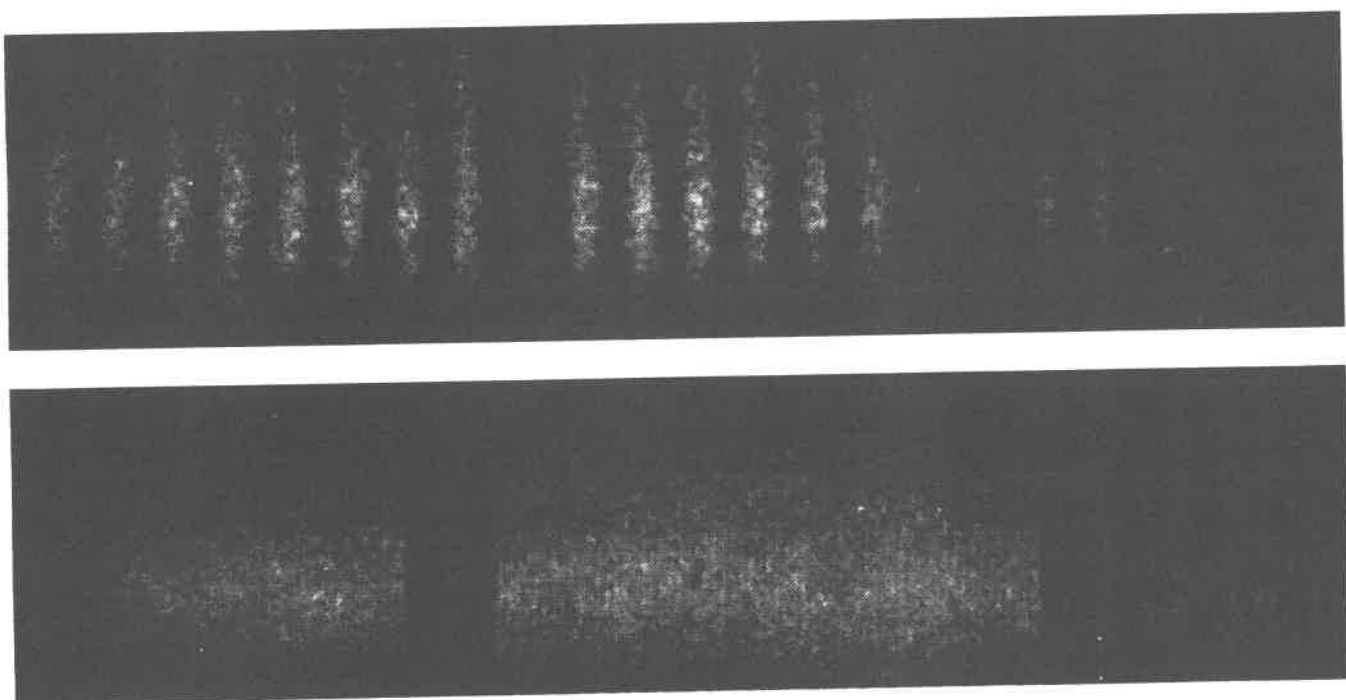


Figure 3: Fringe images at twists of about 0.4 and 2.2 μrad , with shadows of calibration wires.

Test Results. Figure 3 shows typical images captured by computer, including shadows of a pair of steel wires from a variant of the calibration system described above. In the present context the demonstration of stable, high-contrast fringes is sufficient to show that the system is in principle capable of acting as an angle calibrator. These images lose intensity towards their sides because some of the input beam conditioning has been removed for convenience during basic testing when regular major changes are being made to the system. Wider and more uniform images will be available with the full system, to the benefit of the fringe analysis statistics.

The smallest fringes yet resolved in practice have a pitch of around 60 μm , corresponding to an angle of 3.5 μrad (0.7 arcseconds). A similar camera having an 18 mm diagonal could detect fringes down to 90 μm ($\sim 2.2 \mu\text{rad}$). Since fringe spacing is inversely proportional to angle, these limits give the maximum range of the configurations. They overlap usefully with the lower limits of more

conventional methods. Greater range can be obtained by means of either much more expensive cameras or (at a considerable cost of time) photographic films, where fringes as small as $5\text{--}10\text{ }\mu\text{m}$ can be readily detected and $1\text{ }\mu\text{m}$ is feasible ($\sim 200\text{ }\mu\text{rad}$). Smaller angles give wider fringes, so resolution of angle rests with the image analysis technique. The field is about 6 mm wide and there are practical reasons for not much increasing this value, so if one complete fringe is needed within the field the resolution will be about 70 nrad (0.014 arcseconds). Many standard algorithms need several fringes in the field for good reliability, but some will work with sparse data and new developments involving phase stepping [1] should allow good estimation of the pitch even when only a small part of a fringe is in view. In this respect, carefully positioned phase plates can provide one image 'sheared' into three phase steps, figure 4, thus avoiding errors from thermal drift between sequential phase steps, each of which takes about 100 seconds.

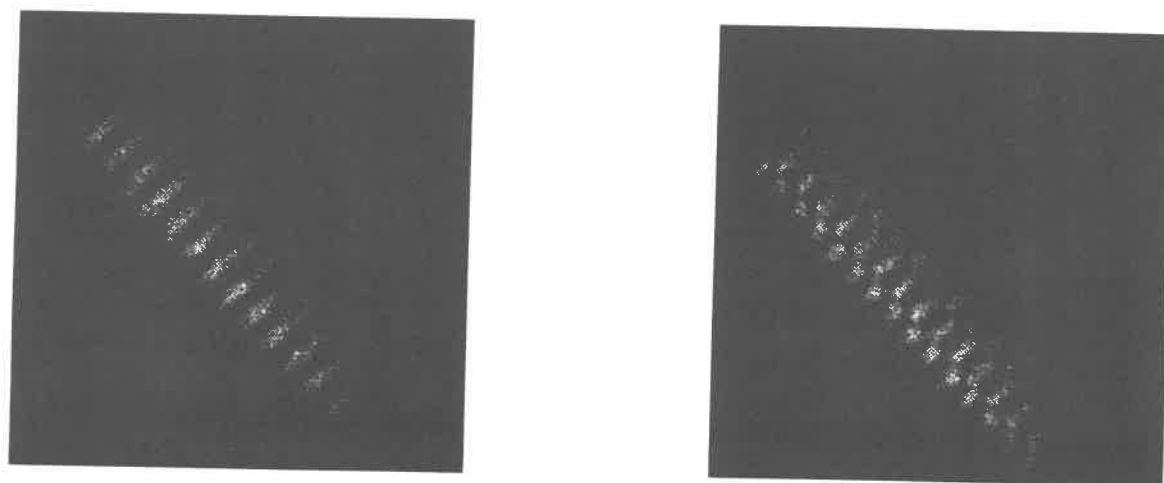


Figure 4: Moiré fringe image: original and simultaneous three-phased-stepped version.

Tests with the x-ray beam in both vertical and horizontal orientations detected no difference in the quality of the data. As contrast of the quality shown was maintained over the full test range in both cases, we conclude there are no practical problems in offsetting the flexure from its ideal (Abbe principle) position. This is advantageous in two respects. First it shows that simple, robust monoliths such as the present one may be used, with no need to introduce more complex mechanisms to project the rotation axis onto the centre of the blade. Second, with the vertical planes, displacement fringes from extremely small angular changes cause large phase shifts in the moiré fringes with virtually imperceptible change of their pitch and may be used for interpolation or for testing local sensitivity limits. Note that calibration of this phase shift depends on knowing the Abbe offset distance: for the present monolith 1 nrad rotation causes about 60° phase shift.

A 10 g laboratory weight placed on the monolith torque arm, directly above the position of one magnet, 30 mm from the centre caused fringes at the limit of image resolution. Thus this torque twists the blade about $3.6\text{ }\mu\text{rad}$, indicating a hinge stiffness approaching 700 N.m rad^{-1} . Given heavy etching after machining to nominal size and the cube-law dependence of stiffness on the notch thickness, this is close enough to the design value of 1000 N.m rad^{-1} to provide further confidence that all system parameters are correctly cross-calibrated.

Discussion. We have demonstrated that a relatively simple silicon monolith can be used as an angular x-ray interferometer, with moiré patterns captured directly onto a full field imaging system. High contrast fringes are easily obtained and the relative twist between the blades determined from the spacing of the fringes. Attaching mirrors to the monolith for using it to assess autocollimators or other optical measurement systems is straightforward and there is no evidence of contrast reduction caused by their weight inducing strain in the monolith. It is particularly important that the angular interferometer provides a continuous measurement as the angle is changed, making it an ideal device for interpolation, not merely calibration at a few points. With the current device, the range and resolution depend on the pixel resolution and the field of view, which is largely a matter of cost. Various forms of signal processing can be used to obtain a 'best' estimate of the fringe spacing, some better in one region of spacings than others, but no one conventional method looks capable of delivering good results over the whole range of the device capability [7]. These issues will be explored further elsewhere: it appears that a technique based on phase-stepping may give considerable improvements over single image methods, especially at larger spacings. With such methods we expect that angle XRI will take its place alongside its linear brother as a valuable tool for top-level calibration.

Acknowledgment

We are grateful to the Paul Instrument Fund of the Royal Society for their support of this work.

References.

- [1] D R Schwarzenberger, D G Chetwynd & D K Bowen 1989 Phase measurement x-ray interferometry. *J. X-ray Sci. Technol.* 1, 134-142
- [2] G Basile, P Becker, A Bergamin, G Cavagnero, A Franks, K Jackson, U Kütgens, G Mana, E W Palmer, M Stedman, J Stümpel & G Zosi 1993 COXI - a combined optical and x-ray interferometer for high precision dimensional metrology. *International Progress in Precision Engineering* (eds. N Ikawa, S Shimada, T Moriwaki, P A McKeown & R C Spragg) Butterworth Heinemann, 248-259
- [3] D G Chetwynd, N O Krylova & S T Smith 1996 Metrological x-ray interferometry in the micrometre region. *Nanotechnology* 7, 1-12
- [4] D K Bowen, Applications of X-ray interferometry, in *Dynamical X-ray diffraction theory and its applications* (eds A. Authier and B.K. Tanner), Plenum Press, 1996 in press.
- [5] D Windisch & P Becker 1992 Angular measurements with x-ray interferometry. *J. Appl. Cryst.* 25, 377-383
- [6] D G Chetwynd, D R Schwarzenberger & D K Bowen 1990 Two-dimensional x-ray interferometry. *Nanotechnology* 1, 19-26.
- [7] E Okuyama & D G Chetwynd 1995 Investigation of data processing algorithms for a micro-radian angular calibrator using an x-ray interferometer. *Control and Instrument Systems Centre Report #22*, Warwick University

Scanning Probe Microscopes for dimensional length measurements

K. Hasche, O. Jusko, L. Koenders, K. Thiele, G. Wilkening, X. Zhao (◆)

Physikalisch-Technische Bundesanstalt (PTB)
Braunschweig, Germany

1. INTRODUCTION

Scanning probe microscopes (SPM) can be used to obtain three-dimensional topographical data of surfaces with very high resolution. Most SPMs utilize piezoelectric (PZT) ceramic actuators to generate, and measure, the scan movement of the probe relative to the sample. Unfortunately PZTs are known to exhibit strong hysteresis, creep, and thermal drift. Although these measurement error sources can be reduced by software, precise measurements can only be obtained by controlling the scan accurately. This can be realized by coupling position sensors like capacitance transducers to the scanner. Their output signal is free of hysteresis and generally some orders of magnitude more linear than the PZTs. Most approaches use the output signal of the transducers as input to a closed feedback-loop, which controls the scan movement. The linearity of the scan then is determined by the inherent linearity of the capacitance transducers including the electronics. A residual nonlinearity of the order of 0.02% has been achieved [1,2]. In practice however, the linearity often is limited by misalignments, i.e. non-parallelism of the capacitor plates and guiding errors of the scanner [2]. Therefore the main advantage of a scanner with built-in capacitance transducers is the high reproducibility and thus the possibility to calibrate the instrument with low uncertainty. This may be performed by scanning a certified standard and analyzing the data [3]. Yet a lower uncertainty can be realized by calibrating the residual non-linearity of the capacitance transducers by an optical interferometer. Three different approaches for the realization of a measurement SPM - „3d-sample scan“, „2d-sample scan“ and „3d-probe scan“ - have been pursued and are described in the following. These instruments are currently used for the traceable calibration of SPM standards and thickness standards, as well as tools for measurement tasks in nano- and micro technology.

2. „SAMPLE SCAN“ SPM - APPROACH

The first microscope is a modified commercial metrological scanning sample AFM (VERITEKT 3 [4]) with a scan range (x,y,z) of 70 μm , 15 μm , 15 μm . The scan is generated by a 3d flexure hinge device with piezo actuators and integrated capacitance transducers acting in a closed feedback-loop. The resolution (x,y,z) is about 1.2 nm, 0.25 nm and 0.25 nm. The AFM has been equipped with three miniature laser interferometers verifying discrete calibration points in intervals of $\lambda/2$ in the entire measurement volume.

The principle arrangement of the laser interferometers in the AFM is given in fig.1. Primarily, the interferometers serve to calibrate the scales, not only referred to the axes x, y, z in the closer sense, but also for parallels of the axes within the entire measurement volume and to reduce the crosstalk of the three movement axes.

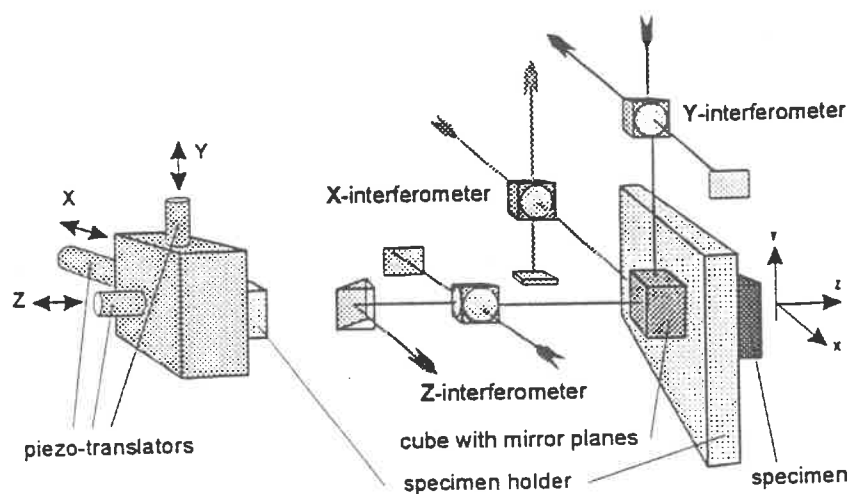


Fig.1 3d-"sample-scan" AFM; principle arrangement of the x-y-z-scanner and of the 3d plane-mirror laser interferometer system integrated in the AFM [4].

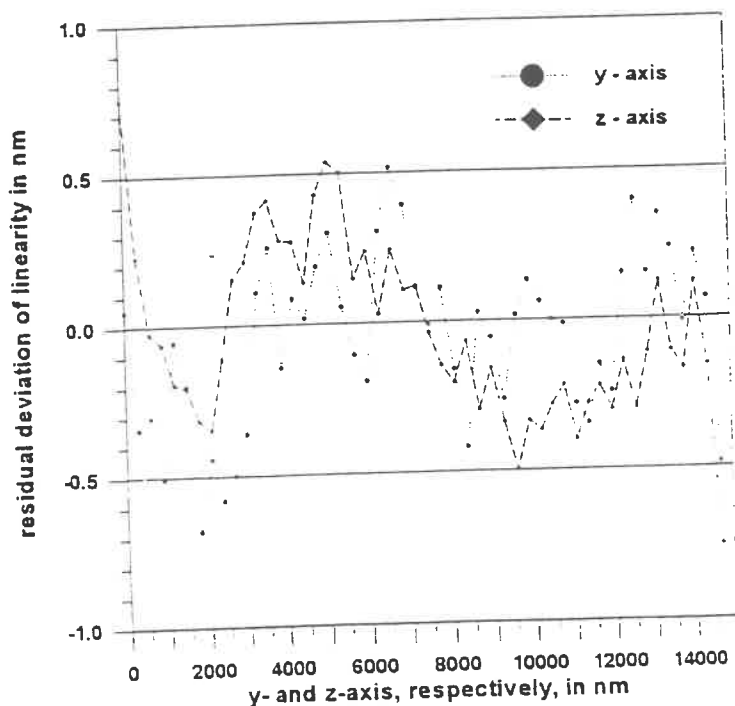


Fig.2 Residual linearity deviations for the x- and y-axis (after compensation)

$$U_{95: 3\text{-dim}} = 5\text{nm} + 2 \times 10^{-4} \times d, d = \text{distance (nm)}.$$

If only one axis is used for the measurement the Abbe offset has no influence. The uncertainty of a step height measurement is estimated to be

$$U_{95: 1\text{-dim}} = 1.1\text{nm} + 2 \times 10^{-4} \times h, h = \text{height (nm)}.$$

The residual deviations of linearity for the y- and z-axis, after (on-line) compensation using the calibration data are illustrated in fig.2.

Fig.3 demonstrates the efficiency of the (on-line) crosstalk compensation. The given example refers to the most critical case of this AFM - the crosstalk from the z-axis to the x-axis caused by the non-orthogonality of the movement axes. The residual displacement in x-direction is plotted as a function of the position on the z-axis for nine

representative positions of x and y, respectively. Abbe errors are caused in the AFM as a consequence of the distance of 23 mm between the cube mirror of the laser interferometers and the x-y-measurement plane. The resulting effects have been determined by angular measurements within the measurement volume and compensated by appropriate lateral displacements in the x-y-planes via the control software. The residual contribution to the uncertainty caused by uncompensated angular deviations is a displacement estimated to be not more than about 5 nm in the x- and y- direction, respectively. The expanded uncertainty of the measurement of any two points within the measurement volume is at present estimated to be

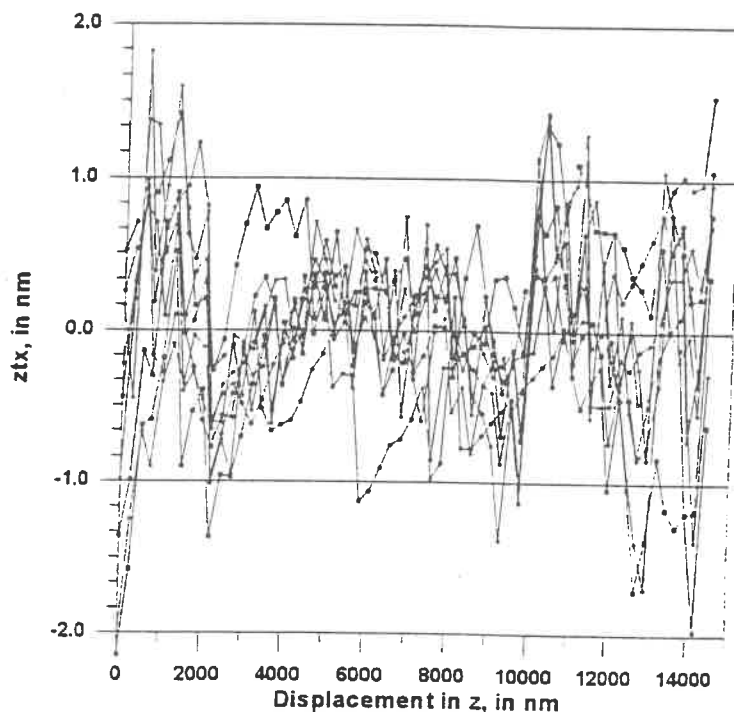


Fig.3 Residual crosstalk z_{tx} from the z -axis to the x -axis for 9 significant positions of x and y .

The second microscope has been built up mostly from commercially available parts [5,6]. The scan range (x,y,z) is $120\ \mu\text{m}$, $120\ \mu\text{m}$, $15\ \mu\text{m}$. It is of the 2d-scanning sample type, consisting of a flexure hinge scanning x - y -table with integrated capacitance transducers and an interchangeable STM or AFM measuring head (see fig.4). The positioning of the x - y -table is controlled in a closed feedback-loop. The STM head consists of a z -piezo actuator with integrated capacitance transducer. The measurement uncertainty for measuring an arbitrary distance d within the measurement volume can be expressed by

$$U_{95, 3\text{-dim}} = 10\ \text{nm} + 3 \times 10^{-4} \times d,$$

d = distance (nm).

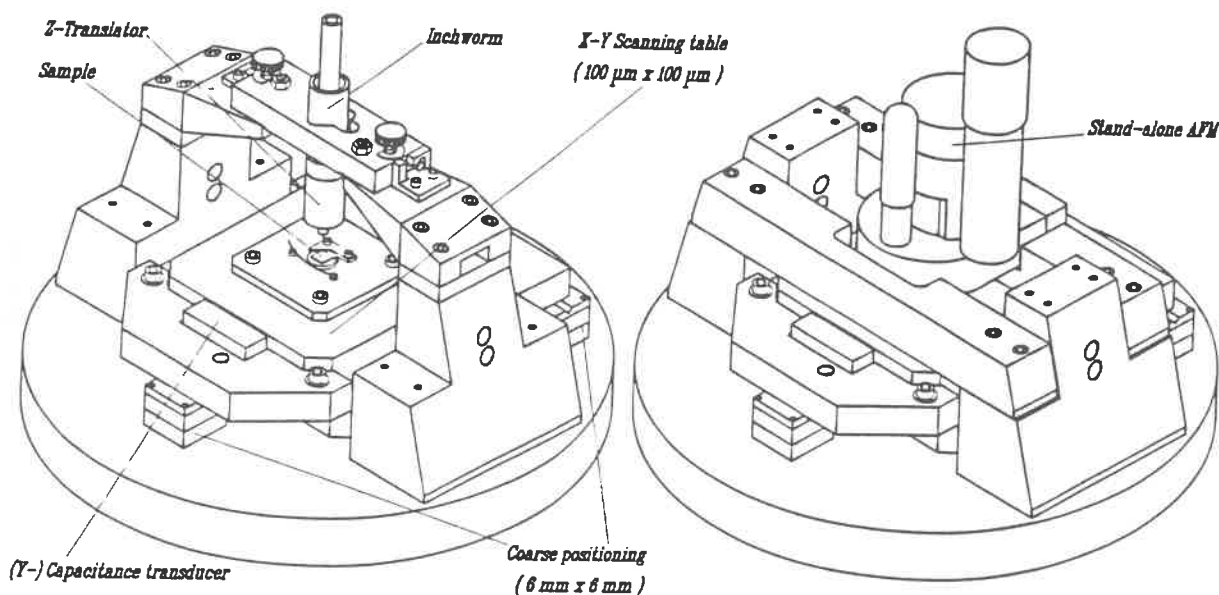


Fig.4 2d-"sample-scan" SPM; STM (left) and AFM version

The calibration is accomplished by using a laser interferometer. The x,y,z axes are calibrated successively by using a miniature prism mirror on the x - y table and a cube mirror replacing the tip as reflectors (see fig.5). This measuring STM has been used successfully for the calibration of standards [7]. However, for practical measurements AFMs are better suited. By replacing the z -scanner of the STM with a commercial stand-alone AFM head [8] the STM can be converted into an AFM-device with a laterally calibrated sample scanner. In fig.4 both possible setups, i.e STM and AFM-mode, are shown. This device may demonstrate that only

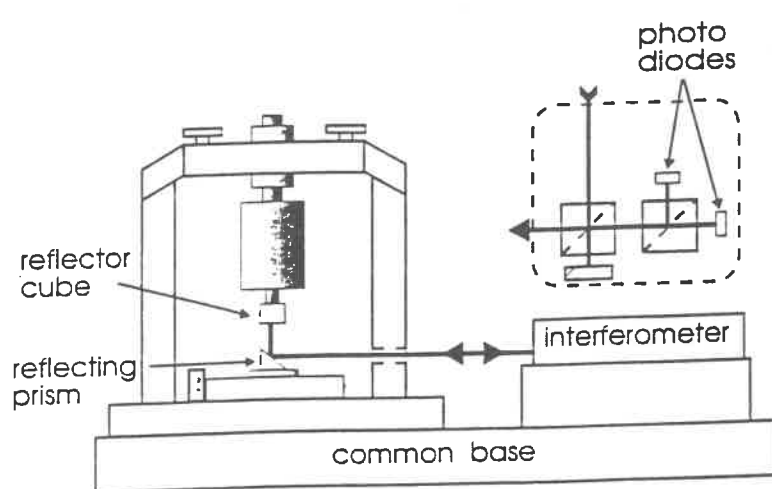


Fig.5 Calibration of the 2d-sample-scan STM; here, the calibration of the z-axis is shown; for x- and y-calibration, the vertical sides of the prism are used.

moderate effort is needed to extend a standard AFM to metrology specifications. In this case an uncertainty can only be given for a single axis measurement. For this case, The uncertainty has been estimated to be

$$U_{95; 1\text{-dim}} = 8 \text{ nm} + 4 \times 10^{-5} \times d,$$

d = distance (nm).

Because the employed AFM lacks an independent gauge for the z-movement, it cannot be used to measure precisely in three dimensions, like for example it is needed for roughness measurements or for calibrating z-standards. Therefore it is necessary to replace the commercial AFM head with a (home-built) head with integrated z-gauging in the future.

3. „PROBE SCAN“ SPM - APPROACH

Measuring SPMs have to be calibrated in the plane in which the actual measurement is carried out (surface of the sample under measurement). This is true for the calibration by standards as well as by interferometry. However, as the sample thicknesses are different in practice the actual measurement plane is likely to differ from the „plane of calibration“. In the case of „sample scan“ SPMs this fact may cause substantial systematic errors. Capacitance transducers, guiding elements of the x-y-table and the sample surface often are situated in different z-heights and the laser beam of the calibrating interferometer again travels in another plane. Such a setup causes a substantial calibration uncertainty by introducing serious Abbe-errors [3,9] in the case of angle errors of the motion of the x-y-table. This problem partly can be overcome by using „probe scan“ SPMs. Here, the plane of calibration and the plane of measurement is always the same, provided that the tip position relative to the x,y,z-gauges stays the same under all circumstances (e.g. tip exchange). It has been shown that an STM working after this principle can be calibrated interferometrically by replacing the tip with a small cube mirror [1,10] and by using a special tip holder with a preadjusted tip [6].

The third microscope is a measuring „probe scan“ SPM consisting of a scanning head with piezo actuators and capacitance transducers in all three axes. Fig.6 shows the design principle of this scanning head.

The x-y-motion is accomplished by means of a piezo bimorph frame. Two parallel bimorphs for each axis form a parallelogram („double leaf spring guidance“). The z-actuator consists of a double piezo plate arrangement. The x-y-actuators and capacitance transducers form a closed feedback-loop. The head is optimized for small Abbe-errors, because the actuators and the capacitors are in-plane, and the angle errors of the scan apparatus are minimized by compensation methods [1]. The scan range (x,y,z) is 30 μm , 30 μm , 6 μm . A calibration uncertainty for an arbitrary displacement measurement (3-dim) has been estimated to be

$$U_{95; 3\text{-dim}} = 8 \text{ nm} + 2 \times 10^{-5} \times d, \quad d = \text{distance (nm)}.$$

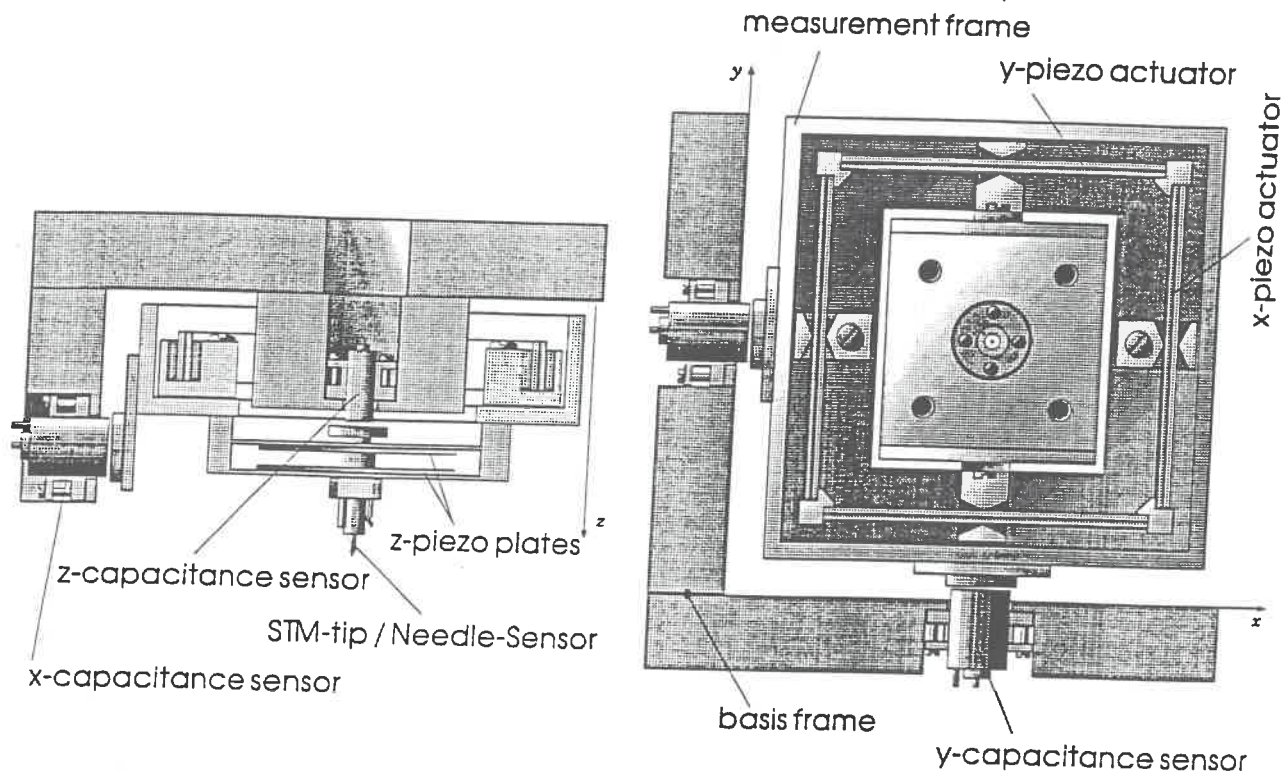


Fig.6 3d "probe scan" SPM head

The conversion of this measuring „probe scan“ STM into an AFM is accomplished by the so-called „Needle Sensor“ [11]. The Needle Sensor may be described as an AFM tip with integrated force sensing (in axial direction). As no external position detection (for force sensing purposes) is needed as it is the case with the conventional AFMs (e.g. triangulation, interferometry) the Needle Sensor is ideally suited for this SPM design. The needle-sensor fits into the already existing tip-holder. In this way STM and AFM operation is easily exchangeable.

4. Conclusions

Three different realizations of metrological SPMs have been shown. All of them use integrated capacitance transducers as position gauges. In all cases the calibration is accomplished by using laser interferometers. From the metrological point of view, the permanent installation of laser interferometers in the SPM for all three axes certainly is the best solution. The interferometers form a separate reference coordinate system and the three axes of the AFM can be investigated and calibrated simultaneously. With the other two SPMs, the axes can be calibrated only successively, and for investigations of the crosstalk and orthogonality, for instance, the built-in capacitance transducers have to be used. The attained uncertainties reflect the metrological (and financial) efforts.

The AFM with built-in laser interferometers rises the question why the interferometers are not used directly for positioning and measuring as they will assure direct traceability. There are mainly two reasons - firstly, the bandwidth of the capacitance transducers is larger than that of interferometers [2], - and secondly, the direct traceability is only true for $\lambda/2$ intervalls as interferometers show non-linearities in the sub- λ -range which are periodic with $\lambda/2$.

The reproducibility of the capacitance transducers in general does not effect the measurement uncertainty. The calibration of the "probe scan" SPM, for instance, showed no change within the stated measurement uncertainty for a period of more than 1.5 years.

Concerning mechanical stability and vibration susceptibility the structures should be as small as possible. The necessity of large scan ranges, sufficient sample space and practical handling led to rather large structures with resonance frequencies of 500 Hz and below. Pneumatic optical tables proved to be sufficient for vibration damping.

The thermal stability is of great importance as due to the large scan ranges the measurement time often is rather long. The employment of material with a low thermal expansion coefficient - as is the case with No.1 AFM - , or high thermal conductivity together with a symmetrical design - as is the case with No.3 SPM - lead to reasonable results. The No.2 STM/AFM showed good results only in a well-stabilized thermal surrounding, as this instrument consists of a mixture of materials due to the fact that commercially available parts have been taken.

5. REFERENCES

- [1] X. Zhao, Hochauflösende dreidimensionale Positionierungsbestimmung bei Rastersondenmikroskopen mittels kapazitiver Aufnehmer; Ph.d. thesis, PTB-Bericht F-20, Braunschweig (1995)
- [2] X. Zhao, G. Wilkening, and K. Horn, High precision three dimensional tip position measurement in Scanning Tunneling Microscopes; Proc. 7th IPES, Kobe, Japan, 937 (1993)
- [3] R. Dürselen, U. Grunewald, and W. Preuß, Präzisions-Piezoscanner für die Nanometrologie; *Feinwerktechnik & Messtechnik*, 483 (1995)
- [4] "VERITEKT 3" is a product of Carl Zeiss Jena GmbH, Germany
- [5] O. Jusko, X. Zhao, and G. Wilkening, A scanning tunneling microscope with high accuracy 3D position control; EOS Topical Meetings Digests Series: Vol 4, ISSN 1167-5357, 41-42 (1994)
- [6] O. Jusko, X. Zhao, H. Wolff, and G. Wilkening, Design and three dimensional Calibration of a Measuring Scanning Tunneling Microscope for metrological Applications; *Rev. Sci. Instr.*, 65 (8) 2514 (1994).
- [7] G. Barbato, K. Carneiro, J. Garnaes, G. Gori, G. Hughes, C.P. Petersen, J.F. Joergensen, O. Jusko, S. Livi, H. McQuoid, L. Nielsen, G.B. Picotto, G. Wilkening; Scanning Tunneling Microscopy Methods for Roughness and Micro Hardness Measurements; Synthesis report on BCR project 3423/1/0/184/4/91-BCR-DK(30), 109 pages (1994)
- [8] TopoMetrix TMX 2000 Explorer is a product of TopoMetrix, Inc., USA
- [9] J. Schneir, T.H. McWaid, J. Alexander, B.P. Wilfly, Design of an atomic force microscope with interferometric position control; *J. Vac. Sci. Technol.* B12(6), 3561 (1994)
- [10] X. Zhao, G. Wilkening, H. Marth, K. Horn; High-accuracy Capacitance Sensor for Positioning of Piezoelectric Displacement Actuator; ACTUATOR 94, June 15-17, Bremen (1994)
- [11] K. Bartzke, and T. Antrack, Atomic Force Microscopic Sensor with Quartz Oscillator and Tip; Proc. UME 3 (International Conference on Ultra Precision in Manufacturing Engineering), ISBN 3-926832-11-8, Aachen, 63-67 (1994)
- "Needle Sensor" is a product of Carl Zeiss Jena GmbH, Germany
- (♦) X. Zhao now is with Physik Instrumente (PI) GmbH, Germany

Toward Accurate Measurements of Pitch, Height, and Width Artifacts with the NIST Calibrated Atomic Force Microscope

R. Dixon¹, T. V. Vorburger¹, P. J. Sullivan¹, V.W. Tsai², T. H. McWaid^{1*}

¹ National Institute of Standards and Technology, Gaithersburg, MD 20899

² Dept. of Material Engineering, University of Maryland, College Park, MD 20742

Introduction

Atomic force microscope (AFM) measurements are being used increasingly for metrological applications such as semiconductor process development and control. Common types of measurements are those of feature spacing (pitch), feature height (or depth), and feature width (critical dimension). If high accuracy is required, the scales of an AFM should be calibrated. Presently available standards for this purpose are calibrated using stylus instruments and optical techniques. The effectiveness of this approach, however, is limited by the differences in the full ranges of the various techniques and by questions of methods divergence. As the repeatability of many commercial instruments continues to improve, the importance of accurate calibration standards will increase. In order to calibrate AFM standards using an AFM, we have designed and built a metrology atomic force microscope, called the calibrated AFM (C-AFM). This instrument has metrology traceable to the wavelength of light for all three axes. To realize this, a flexure translation stage, heterodyne laser interferometers, and a digital-signal-processor based closed-loop feedback system are used to control the x-y scan motion. The z-axis translation is accomplished using a piezoelectric actuator with an integrated capacitance sensor, which is calibrated using a heterodyne laser interferometer. Since the design of this instrument has been discussed in more detail elsewhere [1,2], we will focus here primarily on the results of performance testing of the C-AFM for pitch, height, and width measurements.

Performance Goals and Characterization

At present, our major goals for the C-AFM system are: (1) the capability of calibrating sub-micrometer pitch artifacts with an expanded uncertainty (2σ) of three nanometers [3]; (2) the capability of calibrating sub-micrometer height artifacts with an expanded uncertainty (2σ) of 0.3 nm plus 1% of the height; and (3) the measurement of the top width of sub-micrometer vertical features with an expanded uncertainty (2σ) of 20 nm. To characterize the current level of performance for these measurements, we have compared the C-AFM results with those obtained from other methods. The techniques and instruments used for this comparison include the NIST linescale interferometer [4,5] (used to calibrate standards for optical microscopes), a Talystep[†] stylus profiler (used to perform NIST's step height calibrations), and cross sectional transmission electron microscopy (TEM).

Pitch Measurements

We have characterized the performance of our system for pitch calibrations of artifacts with pitches ranging from 0.2 μm to 20 μm . This was accomplished by comparing the pitch as measured by the C-AFM with the value obtained by using other methods. Comparisons between the NIST linescale interferometer and the C-AFM were made using two samples: a prototype of NIST Standard Reference Material (SRM) 2800 [6], an optical microscope magnification standard, and a prototype of SRM 2090 [7], a scanning electron microscope magnification standard. Lines on these samples with pitches ranging from 4 μm to 20 μm were measured. We also measured, over approximately 50 intervals, the average pitch of an artifact produced by laser focussed atomic deposition [8]. The lines on this sample have a nominal pitch of 212.78 nm, which is traceable to the wavelength of the laser light used in fabrication.

A typical result for this series of experiments is given in figure 1, where results from the measurements of the 10 μm lines on SRM 2090 by both the C-AFM and the linescale interferometer are compared. The regions (1, 2, and 3) on the sample shown as the x-axis correspond to segments of the lines

that are one third of their length (roughly 15 μm) in the transverse direction (perpendicular the measured interval), and the C-AFM values are averages of line profile results over each of these regions. This procedure is necessary because the photodetector of the linescale interferometer averages over these portions of the lines. The error bars given are our estimates of the total measurement uncertainty (1σ) of the C-AFM. The uncertainty of the linescale interferometer measurements on this sample is currently being investigated, but it is estimated to be between 1 nm and 5 nm (1σ). It may be necessary to consider the potential methods' divergence between these two measurement techniques. Since the features on the sample have near-vertical sidewalls, the C-AFM is only sensitive to the edge locations near the top surface of the lines. The linescale interferometer, however, may detect light scattered from a different effective cross section of the feature. This could result in methods' divergence between the two measurements. Although this issue will require further consideration, the important observation is that the residual differences between the C-AFM values and either the nominal pitch or the pitch as measured by the linescale interferometer are consistent within our estimates of the expanded uncertainty (2σ) of the C-AFM measurement.

At the present, we estimate the uncertainty (1σ) for pitch measurements to be 7 nm for large pitch values, on the order of 10 μm . This is believed to be due mainly to motion between the AFM tip and the metrology frame, which realizes the coordinate system for the measurements. Such motion of the tip could result from cantilever flexing or tip deformation. For smaller pitch values, the uncertainty is approximately 3 nm, with the largest contribution being due to the Abbe' offset uncertainty. Other sources of uncertainty are the digital resolution of the interferometer, polarization mixing in the interferometer.

Step Height Measurements

The C-AFM performance on height measurements has been characterized by measuring lithographically produced steps [9][†] ranging in height from about 920 nm down to 18 nm and single-atom Si steps of about 0.3 nm [10]. As recent results of our single-atom step work are being presented elsewhere [11], we will focus here on our measurements of the lithographically produced step height samples.

Currently, the standard uncertainty (1σ) in step height measurements is approximately 1.9% and is limited by variability of the results under different measurement conditions. Some of our results are illustrated in figures 2 and 3 for a nominal 180 nm step. These variations were observed in both contact and intermittent-contact modes of operation. In addition to larger variations in intermittent-contact mode, an offset in the apparent step height of roughly 10 nm was observed between the two modes. Such an offset has not been observed on 18 nm steps. Several effects may be contributing to these observed variations. From other measurements, the calibration of the capacitance gauge itself is not a significant source of variation. Instead, we believe that the response of the force sensor to the upper and lower surfaces can be subject to variations. Cantilever flexing, especially in contact mode, and other cantilever-surface interactions, such as thermal or viscous damping, which are not localized at the tip, are possible explanations for this. In intermittent-contact mode, particularly for larger step heights, differential damping is likely to play an important role in the observed variations. We are continuing to investigate the variations in our measurements. Other possible sources of variation being considered are out-of-plane motions of the translation stage and changes in laboratory temperature and humidity. Smaller sources of uncertainty in the step height measurements also include variability and linearity of the capacitance gauge and Abbe' offset. When we have improved the repeatability of our measurements, we will undertake a detailed intercomparison with measurements using the Talystep[†] stylus profiler which currently performs NIST's step height calibrations.

Linewidth Measurements

As with most commercial AFMs, feature width measurements with the C-AFM are dependent on the geometry of the probe. To obtain an accurate measurement, it is desirable to correct the apparent width of the feature for the width of the probe near its 'apex.' The development of methods and standards for tip characterization is crucial for width metrology, and there has been much recent work in this area [12]. If the geometry of the tip can be estimated, and it is reasonably stable during imaging, it can be deconvolved or eroded from an image to remove some of the probe dilation effects. We are exploring the use of various tip

characterizer samples to estimate the tip size and correct the apparent feature width for the broadening due to the probe. One type of characterizer sample is a cleaved mica surface with colloidal gold particles deposited on it. We have fabricated such samples, using gold particles ranging from 5 nm to 27 nm in diameter, and imaged them in the C-AFM to characterize the probe shape.

By imaging a tip characterizer sample before and after performing measurements on a linewidth sample it is possible to correct the apparent width of a feature for the size of the probe, within some uncertainty dependent upon the severity of the change in the tip from before to after. Figure 4 shows the results from an experiment in which we used this approach to measure the top linewidth (or top width) of preferentially etched lines on a silicon sample. This sample was fabricated by collaborators at the University of New Mexico [13]. After performing the C-AFM measurements, we had cross-sectional transmission electron microscopy (X-TEM) performed on the sample [14][†]. Both the pitch of the lines and the linewidth were extracted from the TEM micrographs. We then scaled the TEM micrograph by equating the pitch from the TEM data to the pitch as measured by the C-AFM. The linewidths were then rescaled by this factor. Severe damage to the C-AFM tip during a width measurement could be readily identified by the extreme changes in the apparent width of the characterizer particles after the top width measurement. When such damage occurred, we typically found larger estimated uncertainties and significant discrepancies between the C-AFM width measurement and the TEM result. If severe damage did not occur, as for measurements 3 and 4, we found good agreement within our estimated uncertainties between the C-AFM width measurements and the TEM result. We have discussed elsewhere [15] more of the details of performing this type of measurement with the C-AFM system and the possible explanations for the variation in our measurements. Although there are some challenges that need to be addressed, we believe that the level of agreement achieved represents a significant step toward accurate linewidth measurements with the C-AFM.

At the present, we estimate that top width measurements of near vertical features using the C-AFM, provided that tip damage is not severe, have an uncertainty of about 15 nm (1σ). We plan to use other types of characterizers and tips to further quantify and reduce the uncertainty of these measurements. Although we believe it should be possible to achieve uncertainties of 10 nm (1σ) or better, we expect that width measurements will continue to be limited by our knowledge of the tip geometry and our ability to correct the apparent feature width for it.

Summary and Conclusions

We have completed a characterization of C-AFM performance for pitch, height, and width measurements, and we are continuing to work on reducing our uncertainties. At the present, we estimate the uncertainty (1σ) for pitch measurements to be 7 nm for large pitch values on the order of 10 μ m and approximately 3 nm for smaller pitch values. The uncertainty (1σ) of our step height measurements of 18 nm and larger steps is approximately 1.9%. At the present, we estimate that top width measurements of near vertical features using the C-AFM, provided that tip damage is not severe, have an uncertainty of about 15 nm (1σ).

Plans for the near future include the testing of different tips and cantilevers to reduce anomalous probe sample interactions and the variations in our measurements. This could be particularly important in reducing our uncertainties in the step height measurements. We also plan to use other types of characterizers and tips to further improve the accuracy of C-AFM width measurements. Longer term goals include the development of a combined pitch/height or three-axis magnification standard for AFMs. We expect such a standard to play an important role in the continued development of AFMs into true metrology instruments.

*Present Address: Tencor Instruments, Santa Clara, CA 95054

[†]Certain commercial equipment is identified in this report in order to describe the experimental procedure adequately. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the equipment identified is necessarily the best available for the purpose.

Acknowledgments

This work was supported largely by the National Semiconductor Metrology Program (NSMP) at NIST. Step height samples were provided by J. Prochazka of VLSI Standards, Inc., and we also thank J. McClelland of NIST and S. H. Zaidi of the University of New Mexico for the provision of samples. Measurements using the NIST linescale interferometer were performed by J. S. Beers and W. B. Penzes. Finally, we give special thanks to C. J. Evans, R. S. Polvani, J. S. Villarrubia, and J. Dagata of NIST for many very helpful and informative discussions.

References

- [1] J. Schneir, T. H. McWaid, R. Dixon, V. W. Tsai, J. S. Villarrubia, E. D. Williams, E. Fu, SPIE **2439**, 401 (1995).
- [2] J. Schneir, T. H. McWaid, J. Alexander, and B. P. Wilfley, *J. Vac. Sci. Technol. B*, **12**(6), 3561 (1994).
- [3] B. N. Taylor and C. E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Tech. Note 1297 (1994).
- [4] J. S. Beers, W. B. Penzes, NISTIR 92-4998 (1992).
- [5] J. S. Beers, NBSIR 87-3625 (1987).
- [6] J. E. Potzick, *Proceedings of SPIE Symposium on Microlithography and Metrology in Micromachining II*, Vol. **2880**-12 (1996).
- [7] B. L. Newell, M. T. Postek, and J. P. van der Ziel, *Proc. SPIE* **2439**, 383 (1995).
- [8] J. J. McClelland, R. E. Scholten, E. C. Palm and R. J. Celotta, *Science*, **262**, 877 (1993).
- [9] Fabricated by VLSI Standards, Incorporated, San Jose, CA 95134-2006.
- [10] E. D. Williams, R. J. Phaneuf, J. Wei, N. C. Bartelt, and T. L. Einstein, *Surf. Sci.*, **294**, 219 (1993).
- [11] V. W. Tsai, T. V. Vorburger, R. Silver, R. Dixon, E. D. Williams, J. Schneir, to be presented at 1996 Annual ASPE meeting.
- [12] J. S. Villarrubia, *Surface Science*, **321**, 287 (1994).
- [13] A. S. Chu, S. H. Zaidi, and S. R. J. Brueck, *Appl. Phys. Lett.*, **63**, 905 (1993).
- [14] TEM analysis performed by R. Anderson of the IBM Analytical Services Group, Hopewell Junction, NY 12533.
- [15] R. Dixon, J. Schneir, T. McWaid, N. Sullivan, V. W. Tsai, S. H. Zaidi, S. R. J. Brueck, SPIE Vol. **2725**, 589 (1996).

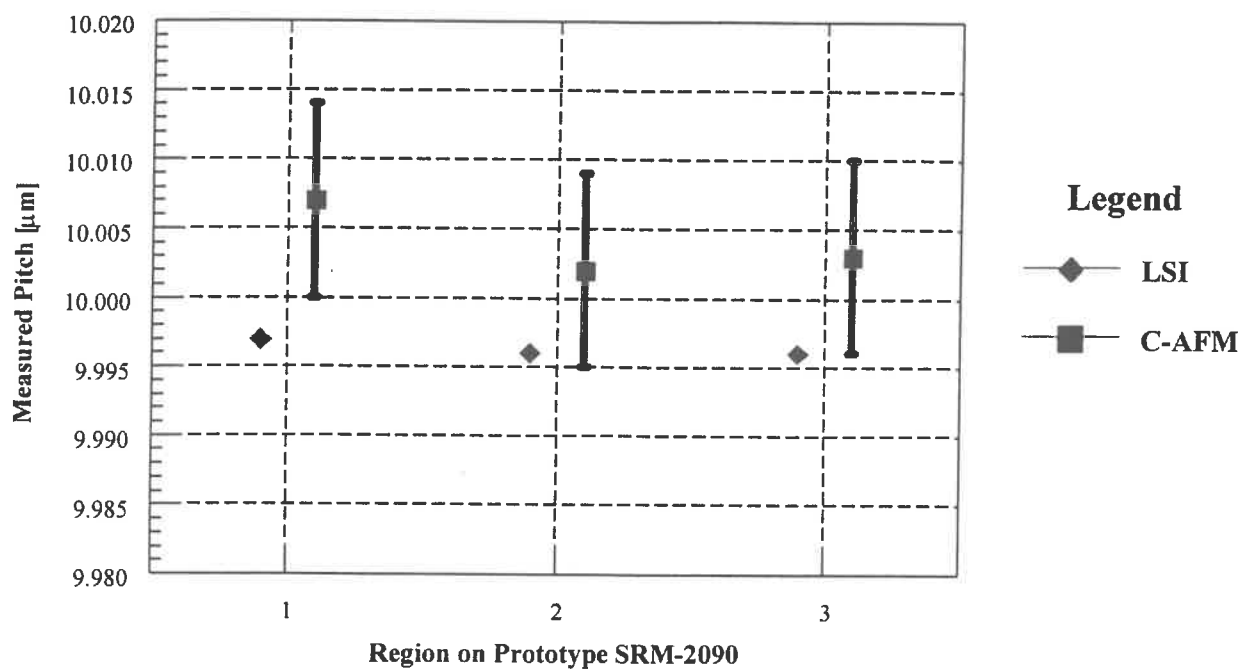


Figure 1. Comparison of C-AFM pitch measurements on prototype of SRM 2090 with results from the NIST linescale interferometer (LSI).

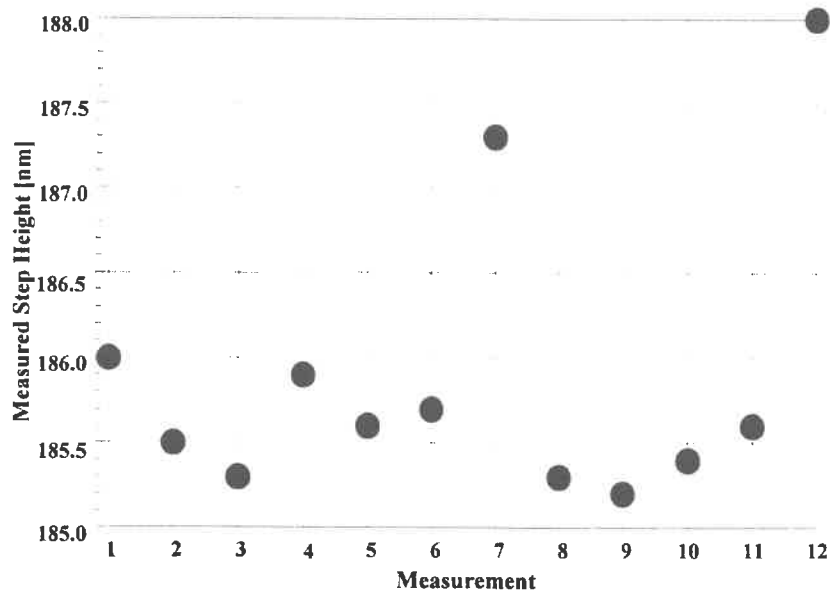


Figure 2. Contact mode measurements of a nominal 180 nm step. Each value represents the average of 10 measurements. The seventh and twelfth measurements were made with different step orientations relative to the fast scan direction.

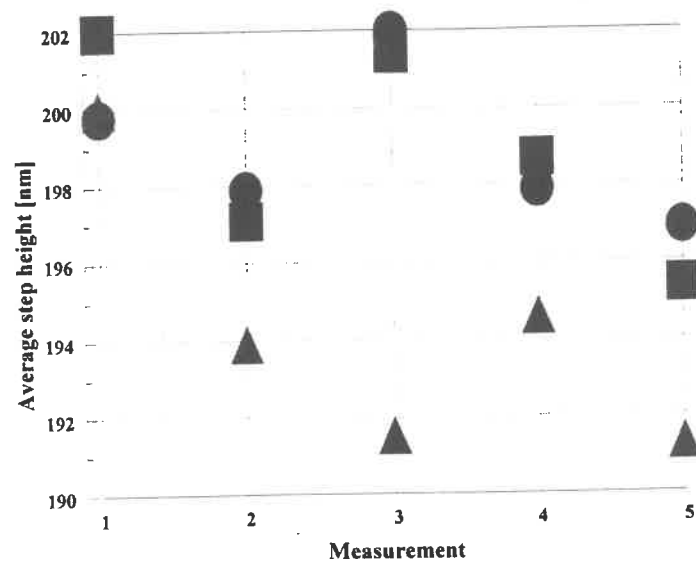


Figure 3. Intermittent contact mode measurements of a nominal 180 nm step. The set point of the z-axis feedback controller was toggled between three values, shown with three separate symbols. These set points correspond to the average cantilever oscillation maintained by the feedback. All data were obtained at the same location, and each value represents the average of 10 measurements.

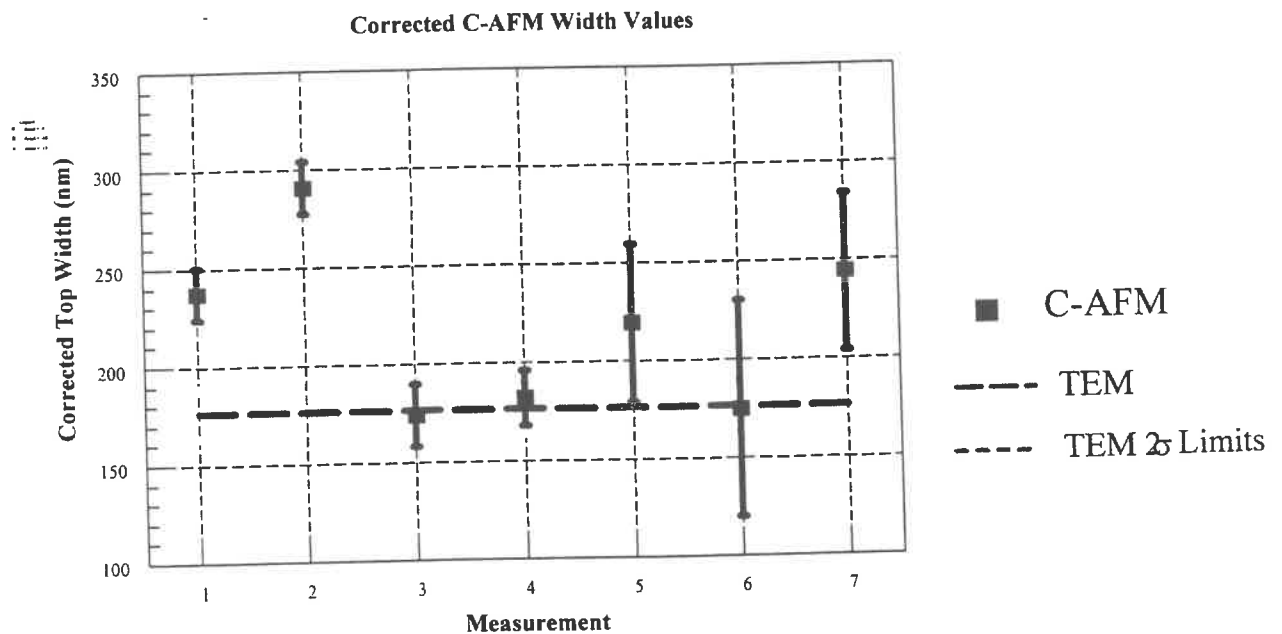


Figure 4. Comparison of C-AFM top width measurements (shown with 1σ error bars) with scale-corrected TEM results (and 2σ limits). When probe damage during the width measurement is not severe, good agreement is found with an estimated 15 nm (1σ) uncertainty.

MEASUREMENTS OF THE SURFACE ROUGHNESS OF GROUND SILICON WAFERS BY CROSS-SECTIONAL TEM

H. Murai, S. Kasaishi, Y. Yokota* and T. Kaneeda

P.O. Hahn**, M. Kerstan** and D. Helmreich**

*Natural Science Research Institute

Okayama University of Science, 1-1, Ridai-cho, Okayama 700, Japan

**Wacker Siltronic, Johannes-Hess-Strasse 24, D-84479 Burghausen Germany

1. Introduction

The TTV (Total Thickness Variation) value of silicon wafer plays an important role in the fabrication of the LSI (Large Scale Integrated circuits)⁽¹⁾. As the integration of the LSI becomes large, the size of each element of the circuit becomes small, and the junction region of n- and p-type becomes shallow. The roughness of the wafer surface and the stacking faults which exist near the surface region (0~0.5 μm depth) are directly affecting the efficiency of producing LSI. For 200mm wafer, the TTV value should be less than the 0.1 μm .

The Si wafer surfaces, mechanically finished by several kinds of wheel grinding, were studied using the cross-sectional TEM (Transmission Electron Microscope) observation.

2. Experiment

The specimen Si wafers were used 200mm in diameter and of {100} orientation. Wafer surfaces were mechanically finished using VPG (Vertical type Precision surface Grinder) and HPG (Horizontal type Precision surface Grinder) with several kinds of wheel grinding⁽²⁾.

The finished conditions were as follows,

1. Single point diamond turning
2. Resin bond wheel grinding #4000 (VPG)
3. Resin bond wheel grinding #3000 (HPG)

These wafers were taken from process line of Wacker Siltronic GMBH⁽³⁾. TEM observation were carried out on a taken from the center part and the quarter of the diameter (D/4) part of the wafer (Figure.1). The cross-sectional specimen were prepared by conventional methods⁽⁴⁾.

3. Discrete Fourier Transformation (DFT)

For the characterization of the wafer surfaces, we applied the DFT (Discrete Fourier Transformation) calculation. The explanation of applying the DFT of wafer surface profile were in shown figure 2~4.

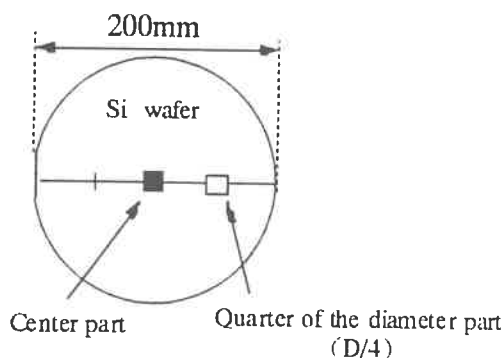


Figure 1. TEM Observation points of the wafer

The figure 2 shows how to measure the wave heights. Tracing the surface shape, the amplitude f of eight points were measured for the same interval.

The equation (1) is the fourier transformation formula. Put $f(t)$ values into the equation (1).

$$G(f) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi ft} dt \quad (1)$$

The exponential part of the fourier term was constructed of sine and cosine functions.

Figure 3 shows fourier analysis of the measured surface profile. The $a_0, a_1, \dots, a_4$ values are the amplitude of sine part. The $b_1, \dots, b_4$ values are amplitude of cosine part.

$$c_n = \sqrt{a_n^2 + b_n^2} \quad (2)$$

The c_n is the square root of power. The a_n and the b_n are the amplitude of sine and cosine part, respectively. Put each values of the a_n and the b_n into the equation (2).

The figure 4 shows the DFT result. X- and Y-axis are the amplitude c_n and spatial frequency, respectively.

The c_0 is average of the amplitude of measured surface. If the c_n values are small for all spatial frequency, the surface is flat. If the c_n values of the low frequency region are high, this means the amplitude of long period waviness is strong and the roughness is large.

4. Results and Discussion

Figure 5 shows the cross-sectional TEM images at the center part of the wafer.

Figure 5 (A) shows the surface finished by diamond tool. The roughness value R_z (the ten point height) was $0.1 \mu\text{m}$. A dislocation depth reached $0.05 \mu\text{m}$ from the surface.

Figure 5 (B) shows the ground surface for VPG #4000. The R_z was $0.02 \mu\text{m}$ and the dislocation depth was $0.13 \mu\text{m}$. Figure 5 (C) shows the ground surface for HPG #3000. The R_z was $0.01 \mu\text{m}$ and the dislocation depth was $0.20 \mu\text{m}$.

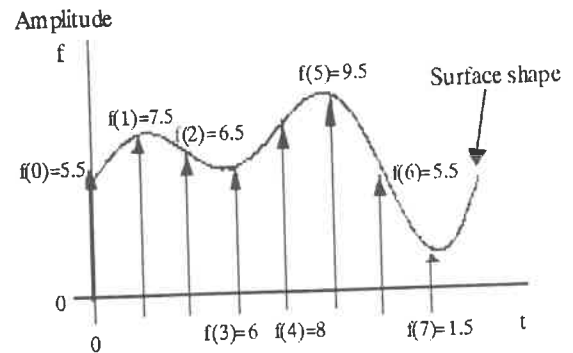


Figure 2. The measurement of the amplitude f

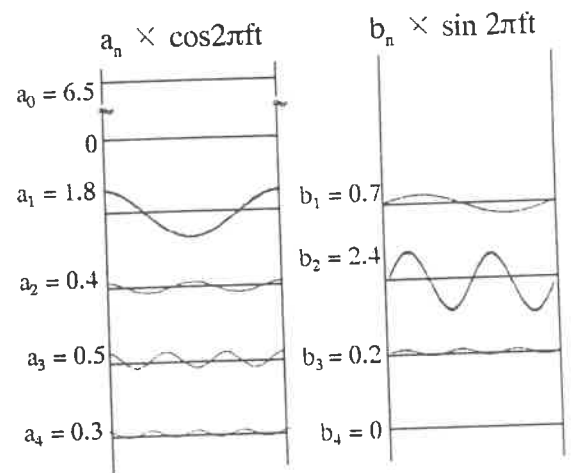


Figure 3. Fourier analysis of measured surface profile.

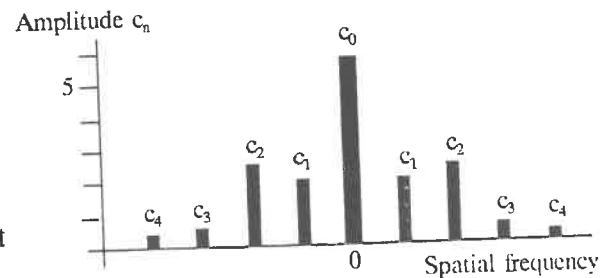
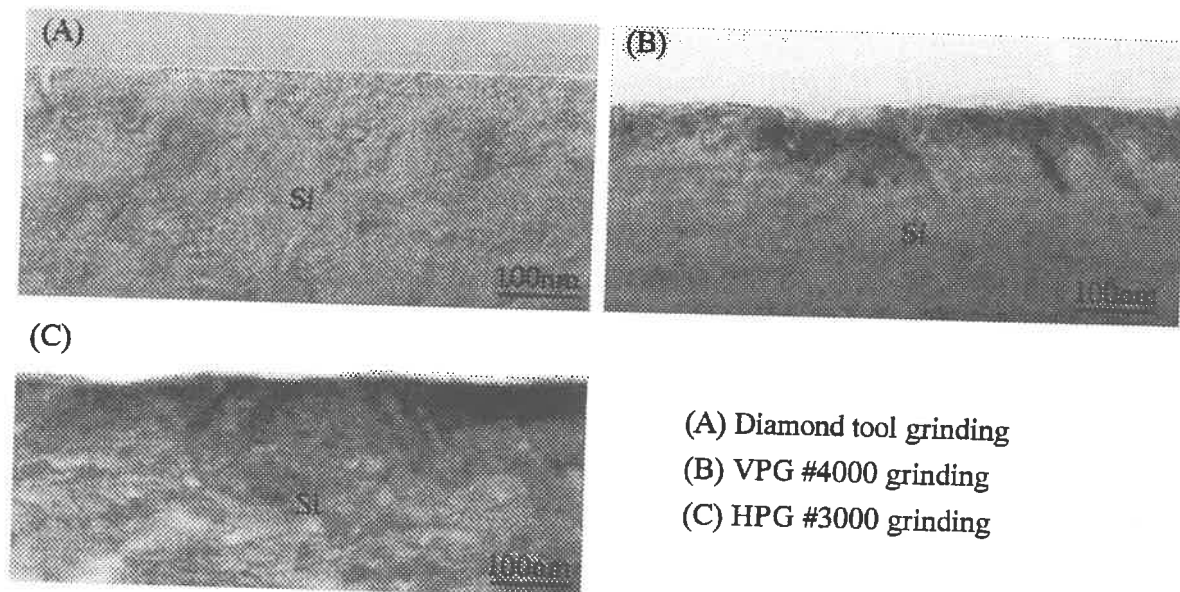


Figure 4. The DFT result of measured surface.

Table 1 shows the surface roughness and dislocation depth measured by TEM micrograph.

Comparing the center part and the D/4 part, the roughness value of D/4 part is larger than that of the center part. Comparing VPG and HPG, the latter has larger dislocation depth. Diamond turning is best among the three for both points.



(A) Diamond tool grinding
(B) VPG #4000 grinding
(C) HPG #3000 grinding

Figure 5. The cross-sectional TEM images at the center part of the wafer

	Dislocation depth	Surface roughness
Diamond tool	0.05 μ m	0.01 μ m
VPG #4000	0.13 μ m	0.02 μ m
HPG #3000	0.20 μ m	0.01 μ m

The center part

	Dislocation depth	Surface roughness
Diamond tool	0.08 μ m	0.01 μ m
VPG #4000	0.16 μ m	0.03 μ m
HPG #3000	0.21 μ m	0.02 μ m

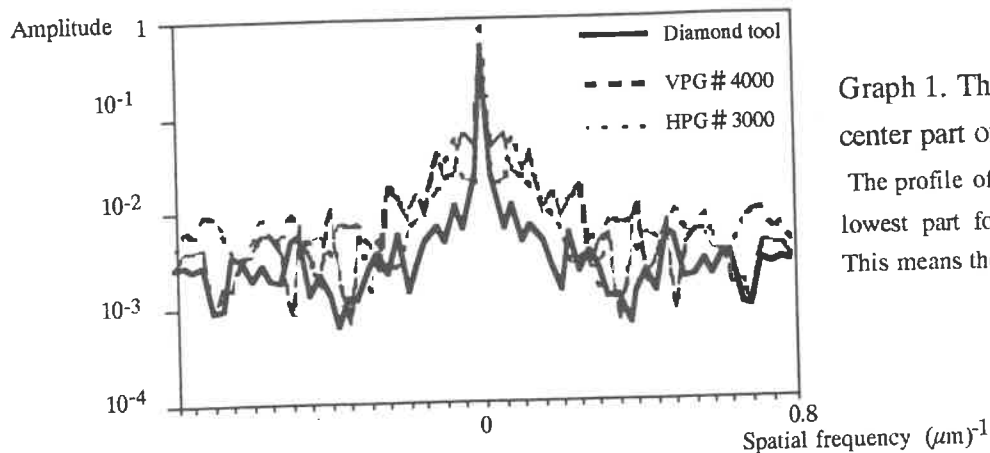
The D/4 part

Table 1. The surface roughness and dislocation depth measured by TEM micrograph

The graph 1 shows the DFT profiles of the surfaces at the center part of the wafer. The surface shapes were traced by cross-sectional images. 64 discrete points were measured for the same interval. Y-axis is the amplitude in log scale and X-axis is spatial frequency.

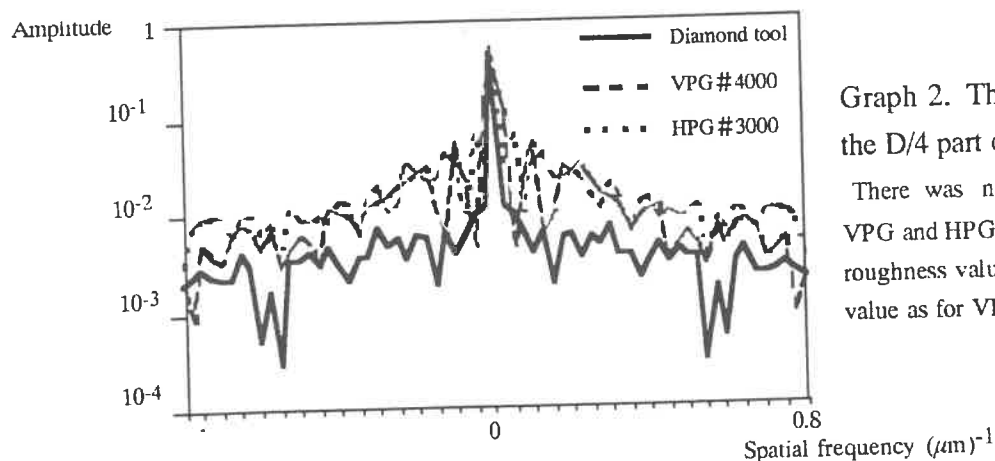
The profile of diamond tool lie in the lowest part for whole frequency region. This means the surface is flat. For VPG #4000, the amplitude of low frequency region is large. This means the amplitude of long period waviness is large and the roughness is large. This means we can measured the surface roughness and waviness by DFT profiles.

The graph 2 shows the DFT profiles of the surface at the D/4part of the wafer. The profile of diamond tool lie in the lowest part of the graph and the surface is flat. There was no difference between VPG grinding and HPG grinding. This means roughness value for HPG is as same value as for VPG.



Graph 1. The DFT profiles at the center part of the wafer.

The profile of diamond tool lie in the lowest part for all frequency regions. This means the surface is flat.



Graph 2. The DFT profiles at the D/4 part of the wafer.

There was no difference between VPG and HPG grinding. This means roughness value for HPG is as same value as for VPG

Summary

The roughness of the nm scale was measured by cross-sectional TEM and the surface shape and dislocation depth could be measured. Comparing the center part and the D/4 part for all wafer, the latter has large both the roughness value and the dislocation depth.

The DFT profile gives the information of not only the roughness but also waviness such as the number of the peak and trough.

Reference

- 1) J. Yoshioka, and F. Hashimoto, M. Miyashita, A. Kanai, T. Abo and M. Daito, "Ultraprecision grinding technology for brittle materials: Application to surface and centerless grinding process", *Grinding Symposium Proceeding, ASME, NY, 1985, p.205*
- 2) T. Abe, Y. Nakazato, M. Daito, A. Kanai and M. Miyashita, "The ductile mode grinding technology applied to silicon wafering process", *Semiconductor silicon 1994*, eds. H.R. Huff et al (*Electrochem. Soc. Pennington 1994*)
- 3) H. Lundt, M. Kerstan, A. Huber and P.O. Hahn "Subsurface damage of abraded silicon wafers" *Semiconductor silicon 1994*, eds. H.R. Huff et al (*Electron. Soc. Pennington 1994*)
- 4) Bravman, J.C. and Sinclair, R. : *J. Electron Microsc. Tech.*, 1,53 (1984)

Fabry-Perot Interferometers for Small Displacement Measurements

Lowell P. Howard, Fred Scire and Jack Stone
NIST Precision Engineering Division, Nano-Scale Metrology Group
Bldg. 220/Rm A-117, Gaithersburg, MD 20899

Introduction

Although both Michelson and Fabry-Perot interferometers are used in instruments designed to detect small displacements, it can be argued that Fabry-Perot interferometers (FPIs) offer some natural advantages due to their multipass nature and their ability to localize an extremely 'narrow' fringe. These perceived advantages are accompanied by a few caveats, however. Alignment is more problematic for multipass interferometers because each pass increases the effective path length and tightens the tolerance window for beam angles. Precise fringe localization in the FPI also means that there is no way to extract phase information from the interferometer cavity in the off-resonance condition because of the absence of a signal. In spite of these conditions, we feel that there are ample rewards to be gained, including traceable displacement measurements down to the atomic and subatomic level.

We are currently 18 months into a multiyear project to realize a compact Fabry-Perot interferometer system for scanned probe microscope z axis metrology and stylus profiler step height calibration. Given such goals, the useful range of our system is expected to be at least 1 μm with a resolution of approximately 1 pm and a displacement measurement combined standard uncertainty of 10 pm. For integration into calibration instruments the system must be located at the distal end of an optical fiber, have a built-in actuator and guiding mechanism and be no thicker than about 25 mm (see Fig. 1.)

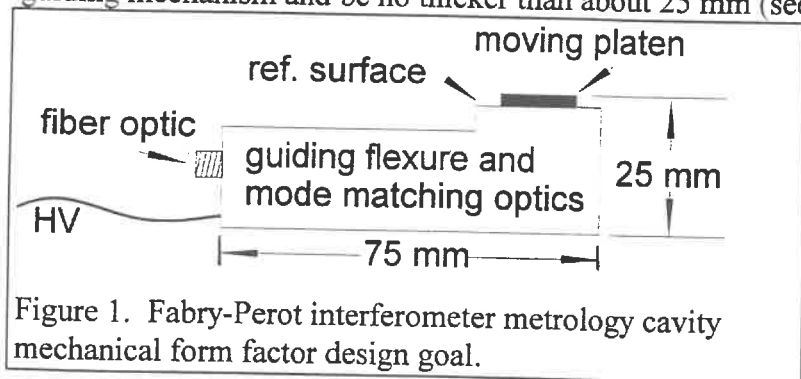


Figure 1. Fabry-Perot interferometer metrology cavity mechanical form factor design goal.

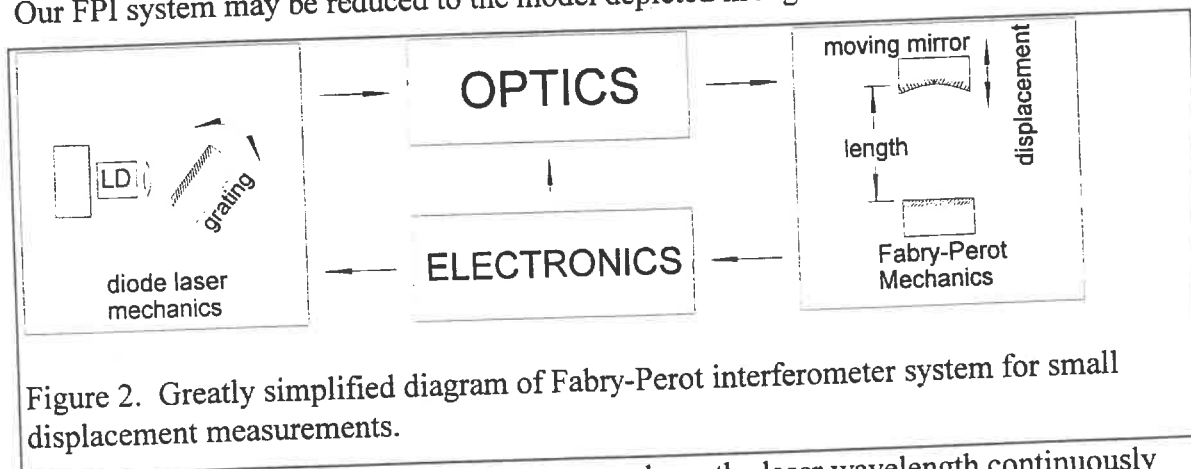
As the use of Fabry-Perot interferometers is somewhat novel amongst traditional metrology circles, a brief survey of some known FPI applications is useful for contrast and comparison. Laser spectroscopy is a traditional domain of the

FPI. Using the FPI for metrology, however, requires access to the interferometer's scanning end mirror and some means to provide a reference plane for alignment purposes. Some workers in the field of x-ray interferometry¹ have used FPIs to make displacement measurements for silicon lattice parameter determinations used in Avagadro constant experiments. Such measurements are, in general, orders of magnitude slower than those used in scanned probe microscopy. Furthermore, FPI use in microscope metrology will not tolerate the gaps in data which are present when the FPI is not tuned to a resonant peak. Dilatometry² for thermal expansion coefficient of materials is another well known use of the FPI. Again, the rates of data acquisition in this application are orders of

magnitude slower than those needed for microscopy. Although the accepted mode of operation of the FPI in this instance, keeping a laser locked to a cavity resonance, avoids the problems associated with having no data in the off resonance condition, the available tracking range with a HeNe laser (≈ 1 GHz) limits the maximum continuously tracked displacement to some 300 nm or less. From even this brief survey, it can be seen that the task of making the FPI useful for microscope metrology will require technical approaches that are some departure from the established norms.

Simplified model of interferometer system

Our FPI system may be reduced to the model depicted in Fig. 2 for discussion purposes.



In operation, the optical and electronic systems keep the laser wavelength continuously locked to a fringe in the FPI. For the locked condition,

$$\Delta F = -\frac{F}{\ell} \Delta \ell, \quad \Delta \ell \ll \ell \quad (1)$$

where ΔF is the change in the locked laser's frequency³, F is the locking laser's nominal frequency, ℓ is the FPI's nominal cavity length and $\Delta \ell$ is the small displacement which is under scrutiny. Substituting in a few typical values for F (474 THz) and for cavity length (50 mm), we see the primary attraction of this method: a sensitivity of $\approx 9.5 \times 10^{15} \text{ Hz} \cdot \text{m}^{-1}$. This high sensitivity means that with only moderate laser frequency control (≈ 10 kHz) we can measure displacements down to the picometer level with interferometric accuracy. Some of the significant limiting factors are: knowledge of the index of refraction inside the FPI cavity, parasitic motions of the FPI mirror, Abbe alignment effects and diffraction phase shifts. Diffraction phase shifts are often more problematic in FPIs than in Michelson interferometers because small (tens of micrometers) beam waists often exist inside spherical FPI cavities.

Referring once again to Fig. 2, we can explore some necessary and useful relationships between the laser cavity behavior and the FPI cavity. The tunable diode laser depicted in the figure contains a diffraction grating as a tuning element⁴; the exact configuration is not important here. What is important to realize is that the motions of the laser tuning mirror are servo locked to the motions of the FPI metrology cavity mirror and comprise essentially a mechanical follower mechanism. Having realized this, the following questions can be asked: what is the optimum relationship between the lengths of the two

cavities? Given this relationship, what role does vibration in the FPI cavity play and what can be said about the relationship of the FPI metrology cavity fundamental mechanical resonance to that of the laser? As to the question of optimum length relationship, we maintain that the laser cavity should be somewhat longer than the metrology FPI cavity based upon the following argument: limiting factors to resolution in both the FPI metrology cavity and the tunable laser will be electronic noise in the actuator drive system and environmental disturbances such as vibrations. Since we do not wish to have *both* the laser and the metrology cavity limited by essentially the same disturbances, we de-sensitize the laser somewhat by increasing the nominal length of its resonator cavity, referring back to eq. 1. In other words, the laser mirror is forced to move through a larger displacement than the FPI metrology cavity mirror. As to the question of vibration in the metrology cavity, we apply a simple first-order lumped-parameter model to the measured prototype interferometer fundamental resonance of 200 Hz. The resulting maximum environmental acceleration would then be $\approx 1.6 \times 10^{-6} \text{ m}\cdot\text{s}^{-2}$ or $\approx 1.6 \times 10^{-7} g$ ($g = 9.8 \text{ m}\cdot\text{s}^{-2}$ is the acceleration due to gravity); this is representative of a very good environment indeed. Assume a cavity length relationship

$$\ell_{\text{laser}} = K\ell_{\text{FPI}} \quad (2)$$

where K is the mechanical 'gain factor' between the two cavity lengths. For a given time interval the laser tuning mirror must move K times as far as the FPI metrology cavity mirror. Therefore the laser mirror must also be moving K times as fast. In a purely mechanical system, the dynamic range-bandwidth product of the laser system must then be K times that of the FPI metrology cavity moving mirror. Given also the fact that the entire system is servo controlled, we must also strive to reduce phase lags in the response of the laser tuning mirror to the displacement of the FPI metrology cavity mirror. For a critically damped second order system, a prudent ratio of fundamental mechanical resonant frequencies might be somewhere around 2, where the laser fundamental mechanical resonance is always higher than that of the FPI metrology cavity. Fortunately, external cavity diode lasers can be electrically tuned small amounts by varying injection current. By using a 'woofer and tweeter' type control scheme where low frequency excursions are mechanically tracked and high frequencies are electrically tuned, phase lags inside the control loop can be kept to a minimum. This example does however, illustrate some of the difficulties encountered when attempting to solve the inverse problem of locking the FPI metrology cavity to a fixed laser source.

Prototype Interferometer

A full, detailed description of the prototype interferometer is beyond the scope of this abstract. The system can, however, be discussed in its block diagram form Fig. 3.

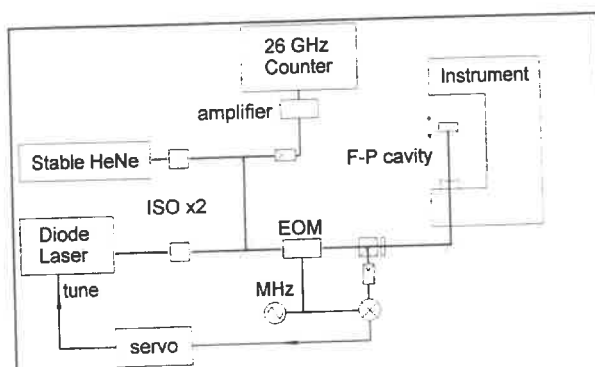


Figure 3. Block diagram of the Fabry-Perot prototype interferometer system.

knowledge of the refractive index of air⁶ at 633 nm. While waiting for commercially produced products, we were ably assisted in our efforts to build a 633 nm diode laser by colleagues from NIST in Boulder, Co. They provided special antireflection coatings on the laser diodes and provided much needed advice on the construction of tunable diode lasers. Our internal effort to produce a laser light source that meets our specifications is depicted in Fig. 4.

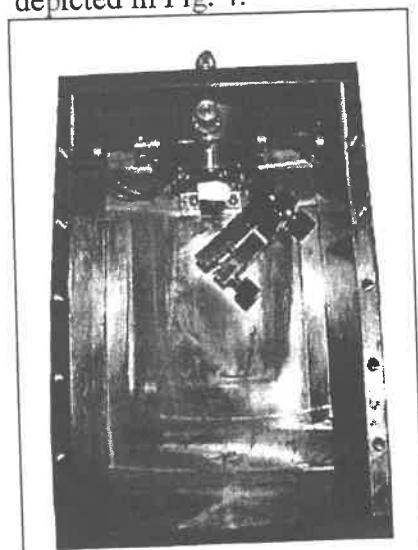


Figure 4. Tunable diode laser cavity on massive copper mounting plate.

The actuator and guiding mechanism for the Fabry-Perot mirrors (Fig. 5.) was constructed by electrical discharge machining (EDM) a solid monolith of invar, selected for thermal stability in an open, testing environment. Subsequent miniature versions of this flexure will be made of Cu or BeCu while attempting to increase the primary structural resonant

The choice of laser wavelength influences many decisions and was complicated by the fact that 18 months ago, when we were picking technologies, 633 nm tunable diode lasers were an immature entity. This is only now being remedied with the introduction of commercially available products. We chose to remain at 633 nm primarily because of the enormous metrology infrastructure already in place with the HeNe laser, most notably the iodine stabilized frequency reference⁵ and the accumulated

The massive copper construction is for isothermal stability. Our laser diode sits upon a thermoelectrically cooled 5 degree-of-freedom mount made also of Cu or BeCu. Since the relationship of the diode facet with respect to the collimating lens was determined to be critical at the sub-micrometer level, a fine adjuster incorporating a demagnification mechanism consisting of a weak spring pushing against a stiff BeCu diaphragm was employed. This mechanism gives us 5.6 μm of laser travel for every 1 mm of micrometer screw travel. A prototype mount of lesser stability has already been used to produce heterodyne beat signals using a polarization stabilized HeNe laser as a frequency reference.

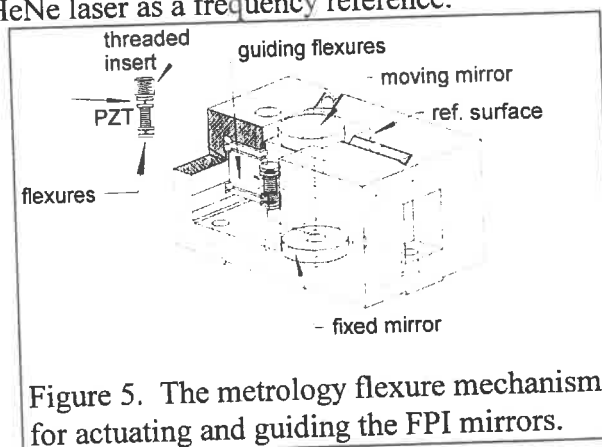


Figure 5. The metrology flexure mechanism for actuating and guiding the FPI mirrors.

frequency, primarily through mass reduction. While invar is far from an ideal material for flexure mechanisms, the required travel of only a few micrometers for a flexure length of ≈ 20 mm make yield strength arguments moot. The scale of the flexure also helps reduce angular guiding errors, measured to be below $0.1''$ of arc using an autocollimator. The flexure is actuated by a low voltage PZT stack decoupled through hourglass flexures located at either end. The actuation point is approximately located at the centroid of the flexure blades for minimizing guiding errors. A reference surface is provided on the top of the invar block along with a set of vee grooves for kinematic mountings. Tungsten carbide inserts line the walls of the grooves for improved hardness and interface stiffness when used with mounting balls. During alignment, a CCD camera is placed on the kinematic mount to aid in making the Fabry-Perot cavity TEM_{00} mode collinear with the axis of flexure carriage travel to reduce Abbe offset.

The fringe locking servo control system functions in a similar fashion to those used for laser frequency stabilization⁷. The servo method, commonly referred to as a 'Pound Drever Hall lock' after some of its inventors, uses an electro-optic modulator (EOM) to introduce radio frequency sidebands to the light impinging upon the Fabry-Perot cavity. Reflected light from the cavity is collected using an optical circulator and processed using two phase sensitive detectors using the EOM driving signal as a reference. Theoretical lineshapes for our cavity conditions are shown in Fig. 6(a) with actual signals shown in Fig. 6(b). The quadrature signal is noteworthy in that a bipolar, linear signal is produced with a zero crossing exactly at the resonance condition. This linear portion of the signal makes an excellent discriminator for a control system scheme. The in-phase signal, with its bipolar spikes at the boundaries of the discriminator signal can be processed and used to validate that the laser is indeed locked to a fringe maximum. The fringe shapes of the prototype interferometer are distorted by at least two factors, both of which are currently being remedied. The extraneous signal present in Fig. 6(b) is presumably caused by nearly degenerate modes present in the current optical cavity configuration. New mirrors should fix this problem. The flattening at the fringe tops is caused by the phase detectors in the signal processing electronics being driven into nonlinear distortion, 'clipping' the signal tops. This too is being fixed and can be expected to yield a much higher fidelity representation of the theoretical signal in Fig. 6(a). Given the conditions of Fig. 6(b), the equivalent electrical noise level is some 3 pm rms for a 50 kHz bandwidth with a sensitivity $\approx 19 \text{ pm} \cdot \text{V}^{-1}$.

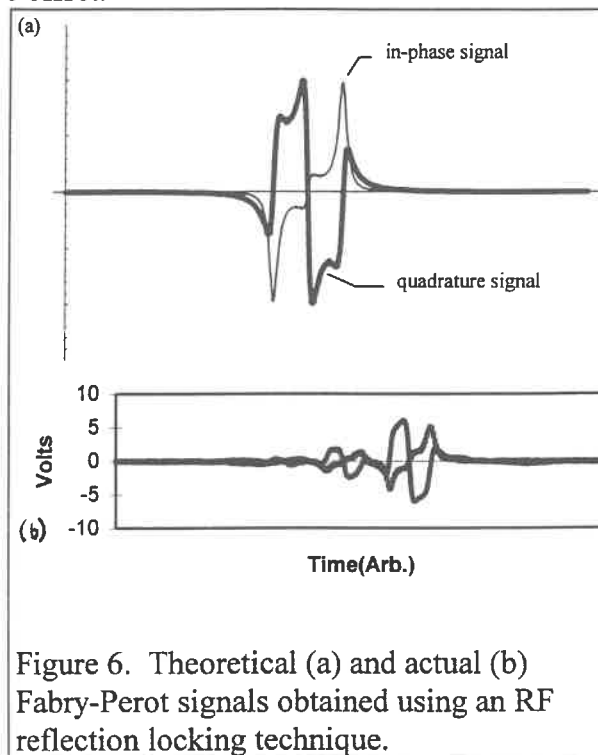


Figure 6. Theoretical (a) and actual (b) Fabry-Perot signals obtained using an RF reflection locking technique.

The project is currently in a state of final assembly. All the major components have been obtained. A significant portion of the systems had to be designed and fabricated in-house. The remaining task of locking our redesigned diode laser to our FPI metrology cavity and measuring the frequency shifts on our microwave frequency counter is not expected to be easy work, but should also hold few unpleasant surprises. Once initial testing is complete, fabrication of a miniature FPI metrology cavity will commence with an eye toward integrating the interferometer into stylus and scanned probe microscope calibrations.

Acknowledgments

The authors would like to thank Michelle Stephens and Leo Hollberg of the NIST Time and Frequency Division in Boulder, Co. for their contributions to the 633 nm external cavity diode laser.

¹ R.D. Deslattes, *Applied Phy. Lett.*, **15**(11), p. 386, (1969).

² S.R. Patterson, *LLNL Report #UCRL-53787*, Feb. 19, 1988 (NTIS, US Dept. of Commerce, 5285 Port Royal Rd., Springfield, VA 22161).

³ In this paper all frequency values are referenced to vacuum wavelength values using a nominal speed of light $c = 3 \times 10^8 \text{ m}\cdot\text{s}^{-1}$. In practice, knowledge of the speed of light in air is a significant limiting factor in the accuracy of any interferometer. Frequency and wavelength units are interchanged for convenience and readability.

⁴ K. Liu and M. Littman, *Optics Lett.*, **6**(3), 117, (1981).

⁵ T.J. Quinn, "Mise en Pratique...", *Metrologia*, **30**(5), 523, (1992).

⁶ P. Ciddor, *Applied Optics*, **35**(9), 1566, (1996).

⁷ R.W. Drever, J.L. Hall, F.V. Kowalski, J. Hough, G.M. Ford, A.J. Munley and H. Ward, *Appl. Phys. B*, **31**, 97, (1983).

A metrological three-axis translator and its application for constant force profilometry

V.G. BADAMI, S. T. SMITH, S.R. PATTERSON

Department of Mechanical Engineering and Engineering Science
University of North Carolina at Charlotte, Charlotte, NC 28223

Introduction

Stylus and constant force transducers form the basis of a broad range of instruments for metrological characterization of fine surface features.^{1,2,3} Typically, stylus based instruments consist of a long range traverse mechanism for translation of the specimen relative to the probe, with the probe system mounted with its axis of measurement perpendicular to the traverse direction. In practice, it is the Abbe errors associated with the misalignment between the axis of the probe that interrogates the surface and the measurement axis of the transducer used to sense the displacement of the probe that dominate the achievable precision in profile height measurement. Currently, a force probe system is being developed to measure fine optical surfaces with an accuracy of 1 nm and a resolution of 0.1 nm for vertical surface profile deviations of up to 3 mm. The heart of this profiler consists of a constant force probe similar to that developed by Howard and Smith,^{4,5} combined with a laser interferometer. Since the Abbe error is due to both the offset and the angle of tilt, control of either parameter can be used to minimize this error. The precision of alignment between the axis of the laser interferometer and the stylus probe is limited by manufacturing considerations. The strategy used to minimize this error is to control the angle of tilt. To this end a novel, three-axis flexure arrangement has been devised that allows for active control of two rotations over a range of 30 μ rad while allowing a displacement of approximately 1.5 μ m. The tilt control mechanism may also be exploited to obtain an estimate of the Abbe offset by monitoring the displacements as measured by the laser and the probe while rocking the probe head through a known angle.

Principle of Operation

A schematic diagram indicating the relationship between the various parts of the device is shown in Figure 1. The force probe is mounted to the target mirror assembly which, in turn serves as a plane mirror for a remote interferometer. The mirror assembly also functions as a target for capacitance gages. The PZT actuators can be controlled individually in combination to provide tilt and translation of the force probe relative to the top plate. Since tilt in two axes is required, a minimum of three actuators and three capacitance gages are required.

The force probe assembly used to interrogate the surface is shown in Figure 2. It consists of a stylus probe mounted on a glass cantilever. Electrodes deposited on the back of the cantilever and the top of the base form the plates of a capacitor. Deflections of the cantilever change the spacing between these plates, changing the capacitance, thus enabling the deflections of the cantilever to be monitored. This design is based on the work of Howard and Smith and is capable of sub-nanometer resolution over a 1 μ m range.

In addition to providing tilt cancellation, the probe head also serves as a high-bandwidth actuator which is used to provide constant force stylus measurement. This is achieved by servo controlling the extension of the probe head to maintain constant cantilever deflection. The probe head in turn is mounted to a low-bandwidth, long-range vertical slide which provides the travel required by the large vertical profile deviations envisaged in current applications. For absolute measurements, the position and tilt of the force probe is measured relative to a metrology frame by interferometric methods. For this purpose, the rear surface of the probe is an optically flat reflector.

The force probe assembly is mounted to the probe head which consists of the actuators, flexures and capacitance displacement sensors. The flexure consists of two sets of three flexure elements (represented schematically in Figure 3 by the springs). The lower set of flexures provide the compliance for the PZT actuators to enable the required displacement and tilt. The upper set is used to adjust the preload on the actuators during assembly and allow for slight dimensional mismatch between the length of the flexure and that of the actuators. The adjustment

screws control the deflections of the more compliant upper flexures which in turn control the preload on the actuators due to the deflection of the lower flexures. Once the preload is set, the controlling stiffness is that of the lower flexures. The deflections of the individual PZT actuators are monitored by capacitance gages which can be used in combination to measure both displacement and tilt of the probe. Although the position of the force probe is monitored by a remote interferometer, the capacitance gages provide a more tightly coupled metrology loop that enhances the controllability of the entire system. In addition, the present design allows for the operation of the device as a stand alone device and facilitates bench testing.

Figure 4 shows an exploded view of the complete assembly. The two sets of flexures are part of a monolithic structure made out of aluminum which also includes slots for the PZT actuators. The flexures are equispaced around the circumference of a circle giving the monolithic flexure three-fold symmetry. Three PZT actuator segments are mounted on the top plate and pass through slots in the flexure. The actuators are made from a single piece of ceramic and exploit the Poisson effect to produce the desired fine motion. These actuators bear on the bottom plate via spherical contact surfaces that allow for the angular tilt. Capacitance gages mounted in the top plate monitor the displacement of each PZT with the bottom plate forming the target for the capacitance gages. The force probe is mounted to the bottom of the target mirror, which in turn is clamped to the lower surface of the bottom cap using a semi-kinematic clamping arrangement (not shown). Changes in length of the PZT actuators cause the bottom plate to tilt and translate relative to the top plate, the latter of which has a large aperture to allow the laser beams of a remote interferometer to pass through to the target mirror. While the aluminum structure of the probe head is susceptible to thermal changes, the force probe assembly and target mirror are constructed out of Zerodur™ and are directly referenced to the metrology frame through the interferometer, thus minimizing the thermal sensitivity of the main measurement loop. Monolithic construction has been employed to minimize joint losses and hysteresis and to obtain a compact, high stiffness structure.

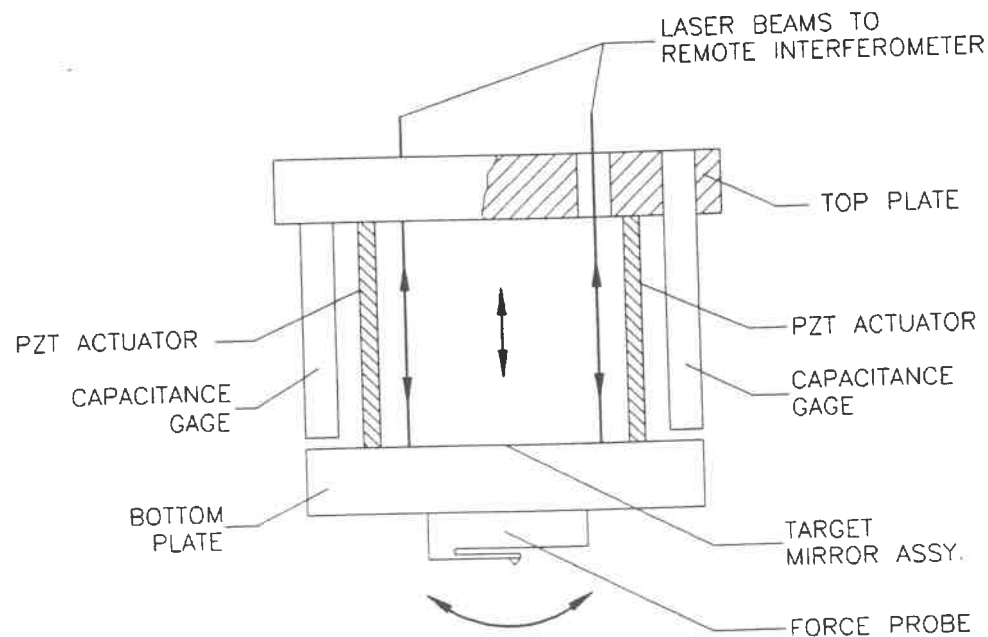


Figure 1: Relative dispositions of capacitance gages, interferometer and PZTs

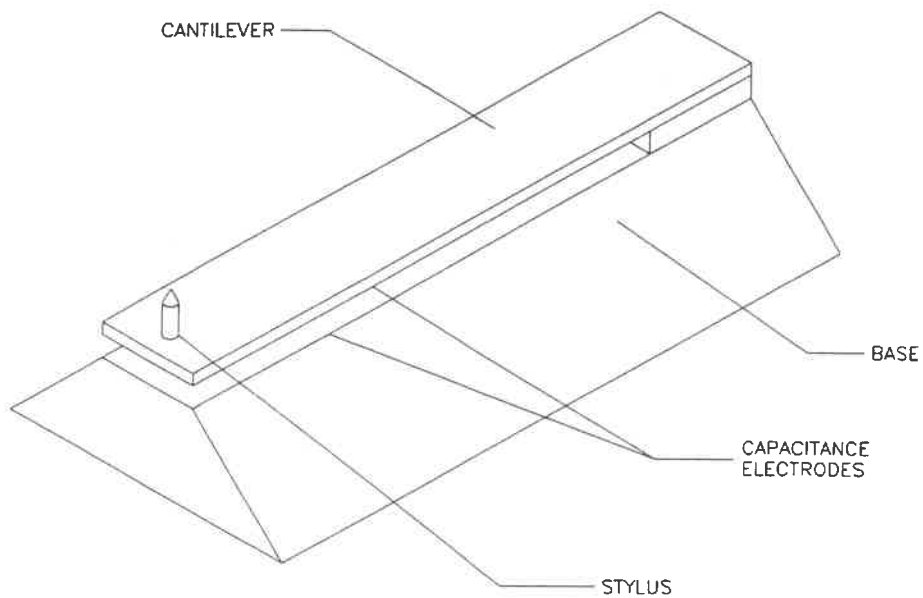


Figure 2: Schematic of force probe

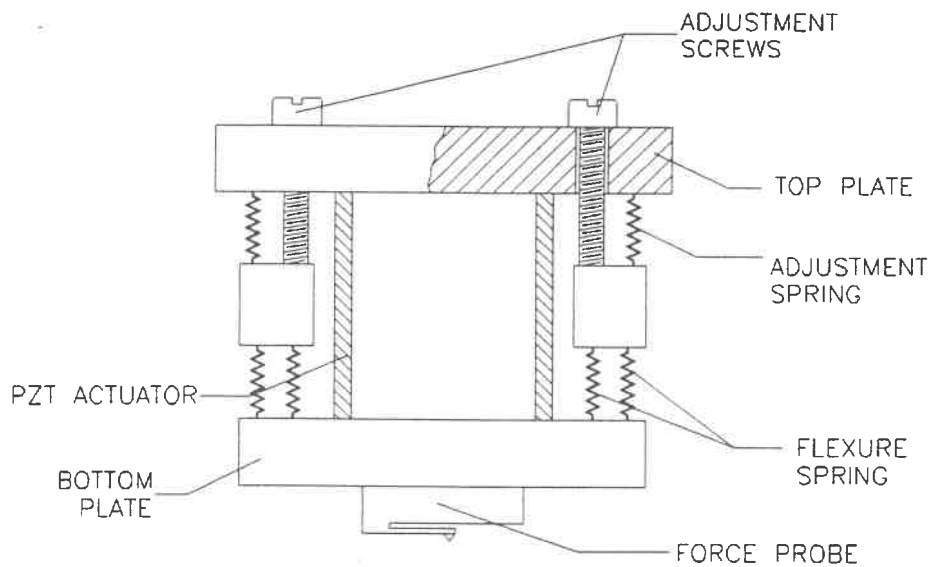


Figure 3: Schematic of flexure arrangement

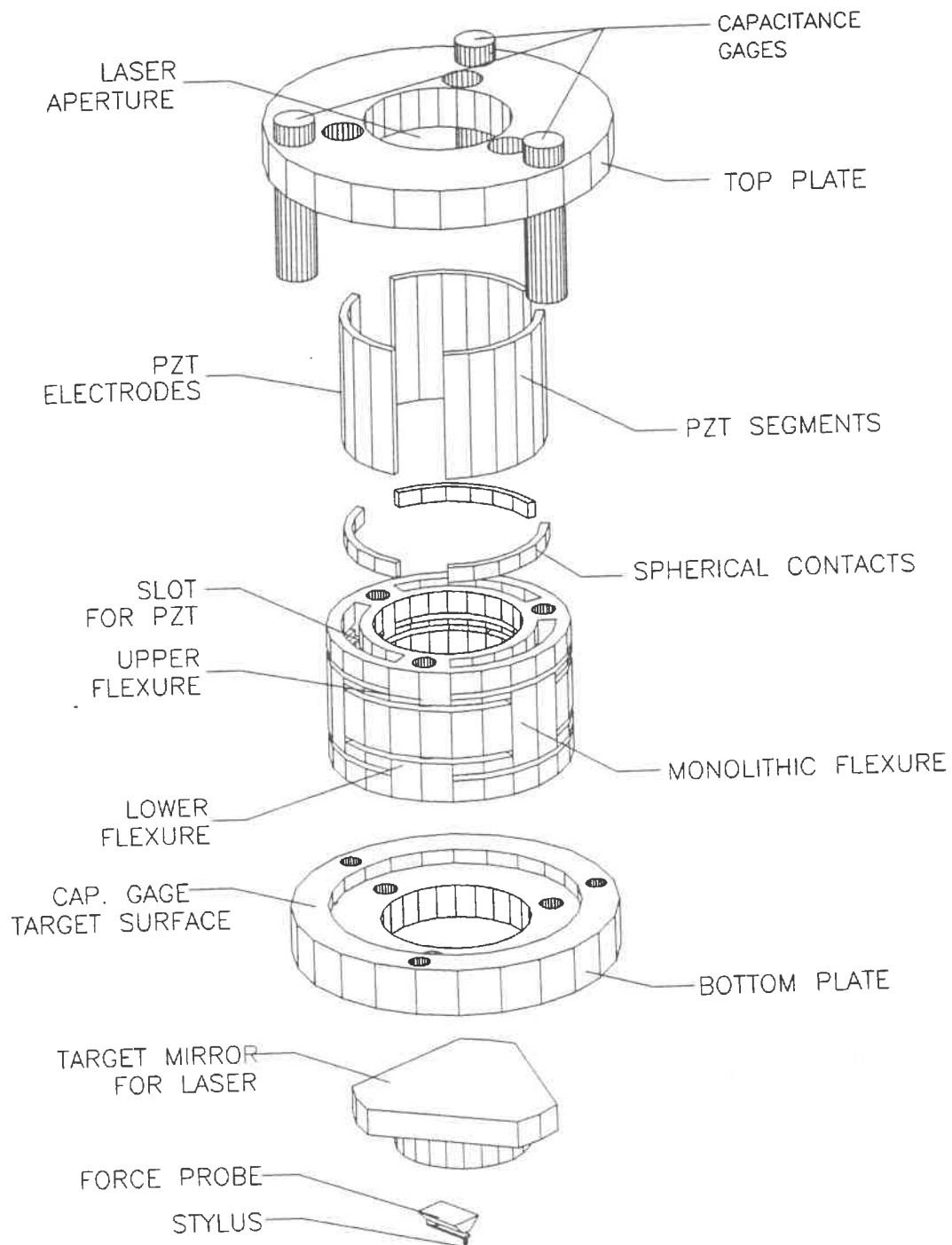


Figure 4: Exploded view of assembly

References

- ¹ Thomas, T. R., Rough Surfaces, Longman Press, London, 1982.
- ² Chetwynd, D. G., and Smith, S. T., "High precision surface profilometry: From stylus to STM," *From Instruments to Nanotechnology*, ed. Gardner, J. W., and Hingle, H. T., Gordon and Breach, London, 1991, pp.273-300.
- ³ Sarid, D., Scanning force microscopy: with applications to electric, magnetic and atomic forces, Oxford University Press, NY, 1994.
- ⁴ Howard, L. P., and Smith, S. T., "A metrological constant force profiler," *Rev. Sci. Instrum.*, 65(4), 1994, pp. 892-902.
- ⁵ Smith, S. T., and Howard, L. P., "A precision, low force balance and its application to atomic force probe characterization," *Rev. Sci. Instrum.*, 65(4), 1994, pp. 903-903

Modeling and Control of Hysteresis in a Piezoceramic Actuator

Ping Ge and M. Jouaneh

Department of Mechanical Engineering
University of Rhode Island, Kingston, RI 02881

Abstract

The hysteresis nonlinearity and output drift are two major limitations in open loop operation of piezoceramic actuators used in ultra-precision displacement applications. For tracking control applications, the hysteresis seriously affects the tracking performance even with the use of closed-loop feedback control. This paper presents a new approach for modeling the hysteresis nonlinearity of the actuator under non-cyclic input variations using a modified generalized Preisach model. Based on this model, a tracking control system for a piezoceramic actuator is developed by combining a PID feedback controller with a hysteresis linearizing scheme implemented in a feed-forward loop. Tracking control experiments of arbitrary input signals performed on a stacked piezoceramic actuator showed that a significant improvement of the tracking control accuracy can be obtained by using this tracking control scheme.

Generalized Preisach model for hysteresis

The classical Preisach model can be used to describe many physical phenomena with nonlinear hysteresis behavior [1]. The hysteresis nonlinearity to be presented by this model must satisfy two important properties: the wiping-out property and the congruency property. Our experimental results show that the major hysteresis loop of a piezoceramic actuator is well shaped, and it does not depend on how the loop is approached [2]. This implies that any input extremum with small amplitude does not affect the future output values if the subsequent input value is larger than this input extremum. However, the hysteresis nonlinearity of a piezoceramic actuator does not satisfy the congruency property. This can be seen from Figure 1. When the input voltage starts from u_1 , say $u_1 = 200\text{v}$, and monotonically increase to u_2 , say $u_2 = 300\text{v}$, the output increments from different ascending loops are unequal, $\Delta_1 \neq \Delta_2$ in the figure. Such behavior also holds for the descending loop.

Consequently, when the hysteresis output runs on a minor loop (see Figure 2) as would be the case when tracking arbitrary input signals, a significant distortion will occur if we use the classical Preisach model to predict the hysteresis output of the actuator. To overcome this limitation, the generalized Preisach model needs to be developed for modeling the hysteresis output of a piezoceramic actuator for arbitrary types of input signals.

The generalized Preisach model can be mathematically expressed as

$$x(t) = \iint_{R_{u(t)}} \mu(\alpha, \beta, u(t)) \gamma_{\alpha\beta}[u(t)] d\alpha d\beta + x_{u(t)}^+ \quad (1)$$

where, $\mu(\alpha, \beta, u(t))$ is a weight function of the Preisach model; $x_{u(t)}^+$ is the hysteresis output defined on the major ascending hysteresis loop; $\gamma_{\alpha\beta}[u(t)]$ is the hysteresis operator which has two output values, 0 and 1; α and β are the switching values of input, corresponding to the maximum value of input extremum and the minimum value of input extremum respectively; $R_{u(t)}$ is a rectangular region which can be subdivided into two sets, $S^+(t)$ and $S^-(t)$ (see Figure 3). In $S^+(t)$ the output value of the switching operator, $\gamma_{\alpha\beta}[u(t)]$ is equal to 1 while in $S^-(t)$, $\gamma_{\alpha\beta}[u(t)]$ is equal to 0. In the α - β diagram

of Figure 3, the horizontal interface link corresponds to an increment of the input on an ascending hysteresis loop, thus it always goes up, while the vertical link corresponds to a decrement of the input on a descending hysteresis loop, thus it always moves towards the left. Since $\gamma_{\alpha\beta}[u(t)]$ is equal to 0 in $S^-(t)$, the above Preisach model (1) can be simplified as

$$x(t) = \iint_{S^+(t)} \mu(\alpha, \beta, u(t)) d\alpha d\beta + x_{u(t)}^- \quad (2)$$

To obtain an expression for the weight function, we introduce a new function, $X(\alpha, \beta, u(t))$, which is defined as an output difference at a certain value of input, $u(t)$, between a first-order reversal curve and a second-order reversal curve that is attached to the first-order reversal curve (see Figure 2), i.e.,

$$X(\alpha', \beta', u(t)) = x_{\alpha'u} - x_{\alpha'\beta'u} \quad (3)$$

Substituting for $x_{\alpha'u}$ and $x_{\alpha'\beta'u}$ in Equation (3) by using the facts shown in the Figure 4, gives

$$X(\alpha', \beta', u(t)) = \iint_{R(\alpha', \beta', u)} \mu(\alpha, \beta, u(t)) d\alpha d\beta = \int_{\alpha'}^{\alpha''} \left[\int_{\beta'}^{\beta''} \mu(\alpha, \beta, u(t)) d\beta \right] d\alpha \quad (4)$$

From Equation (4), we can see that the double integral of the weight function over the rectangular region R can be evaluated from first and second order reversal curve output data. This fact is used in obtaining an explicit formula for the generalized Preisach model which can be written as:

$$x(t) = x_{u(t)}^- + \sum_{k=0}^N [X(\alpha'_{k+1}, \beta_k, u(t)) - X(\alpha_k, \beta'_k, u(t))] \quad (5)$$

where, $x_{u(t)}^-$ is the hysteresis output value defined on the major descending loop and $X(\alpha_k, \beta_k, u(t))$ as defined in Equation (3). The details of the derivation of Equation (5) are given in reference [3].

Model-based tracking control scheme

Because of the existence of hysteresis in piezoceramic actuators, the traditional PID feedback controller is inadequate to obtain a good tracking control performance. Hence, we propose a tracking control scheme which consists of a PID feedback control loop augmented with a hysteresis model in a feed-forward loop. In the feed-forward loop, a model-based linearization of the nonlinear hysteresis is performed using the generalized Preisach model developed earlier.

Figure 5 shows the flow chart of the linearization algorithm. The linearizing algorithm involves the following steps: (i) The arbitrary reference signal is discretized and the initial control voltage is set to zero; (ii) the reference displacement is sampled and the hysteresis output of the actuator is computed by using the numerical generalized Preisach model. The calculated output is compared with the corresponding reference displacement, and the error between the reference and actual trajectories is used to compute the updated control voltage. This step is repeated iteratively until a specified accuracy of output displacement is obtained; (iii) The index k is increased and step (ii) is repeated for the next reference displacement value. Here, the new initial value of control voltage

is set to the last updated control voltage. It should be noted that in the linearization process, the control voltage extrema are determined from the reference displacement signal, i.e., when the reference displacement reaches a maximum (minimum) value, the control voltage also arrives at a maximum (minimum) value. This control voltage is, thereby, recorded as the previous input extremum and used in the Preisach model to determine the future hysteresis output.

The proposed tracking control system was designed by adding the above described feed-forward linearizing scheme to a PID feedback control loop. Figure 6 shows this control scheme. In the figure, the control voltage, $u_d(kT)$, is the linearizing term which is selected from a reference displacement/control voltage look-up table. The error signal, $\Delta u(kT)$, derived from the PID feedback controller, functions as the compensation term. As a result, the total control voltage sent to the piezoceramic actuator is the sum of these two signals, i.e.,

$$u(kT) = k_p e(kT) + k_i \sum e(kT) + k_d [e(kT) - e((k-1)T)] + u_d(kT) \quad (6)$$

where, T is the sampling time of the control system.

Experimental results and conclusions

The experimental setup is shown in Figure 7. The control system performance was tested for tracking an arbitrary reference displacement signal. A comparison was made between an open loop control system, an open-loop feedforward linearizing control system, a PID feedback control system and a PID feedback control system augmented with a linearizing scheme in the feed-forward loop. Figure 8 shows the tracking control signals and their errors. The maximum error for the PID feedback controller augmented with a linearizing scheme in a feed-forward loop is about 1/4 of that of an open loop system and 1/2 of that of the regular PID feedback control system. The results show that after incorporating a model-based feed forward loop in a PID feedback control system, the tracking control error was significantly reduced.

References

- [1] Mayergoyz, I., Mathematical Models of Hysteresis, Springer-Verlag, 1991.
- [2] P. Ge and M. Jouaneh, "Modeling hysteresis in piezoceramic actuators", Precision Engineering, 1995, Vol. 17, pp. 211-221.
- [3] P. Ge, "Modeling and control of hysteresis in piezoceramic actuators", Ph.D. dissertation, Dept. of Mechanical Engineering, University of Rhode Island, May 1996.

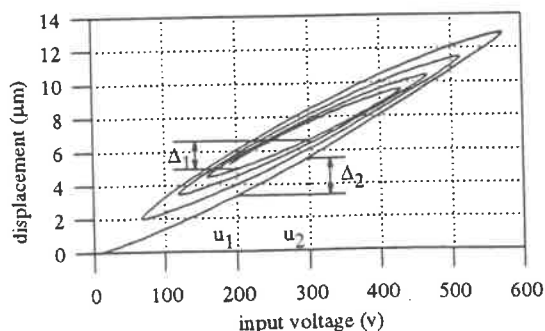


Figure 1. Hysteresis loop of a piezoceramic actuator showing a non-congruent behavior for the same change in input (from u_1 to u_2)

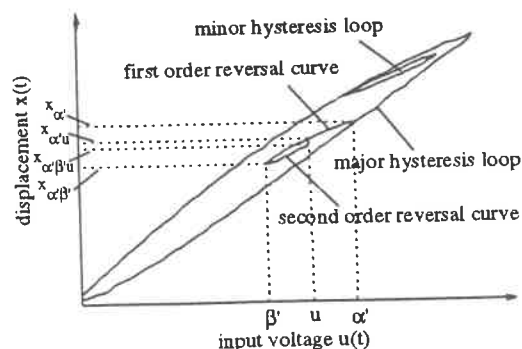


Figure 2. Illustration of the Preisach function $X(\alpha, \beta, u(t))$ defined on first-order and second-order reversal curves

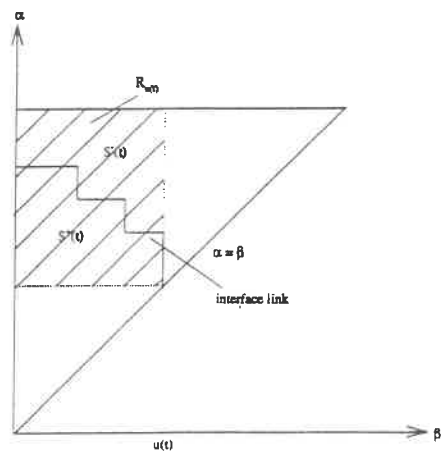


Figure 3. Geometrical interpretation of the generalized Preisach model

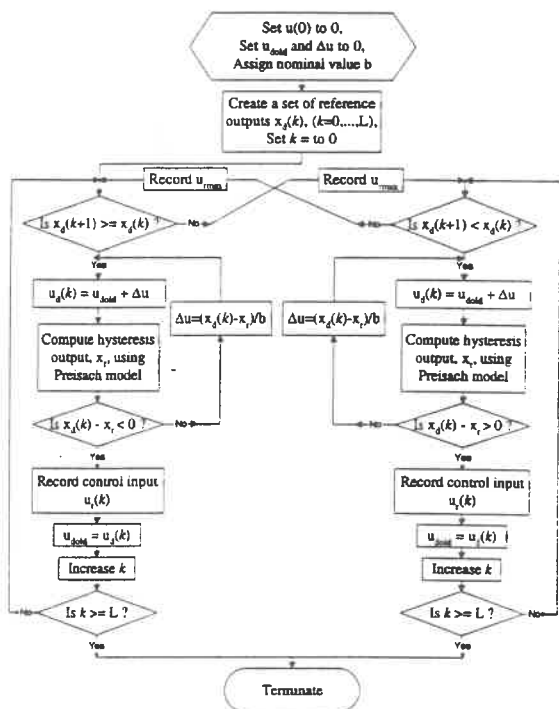


Figure 5. A Model-based linearizing control algorithm

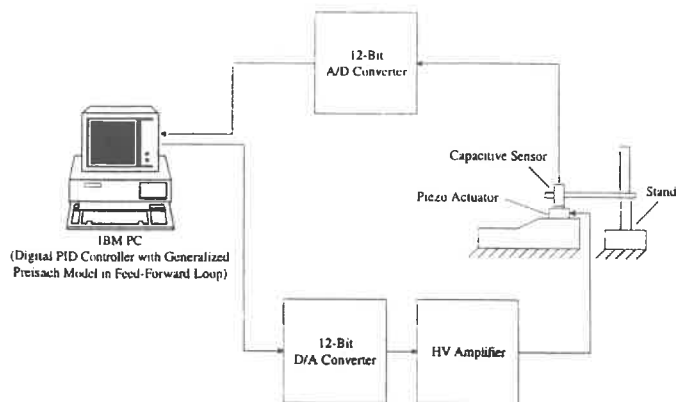


Figure 7. Experimental setup for the piezoceramic actuator control system

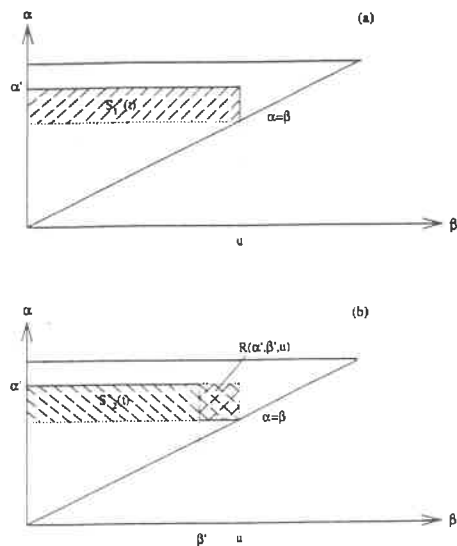


Figure 4. Illustration of the region $R(\alpha', \beta', u)$ used for computation of the function $X(\alpha', \beta', u)$

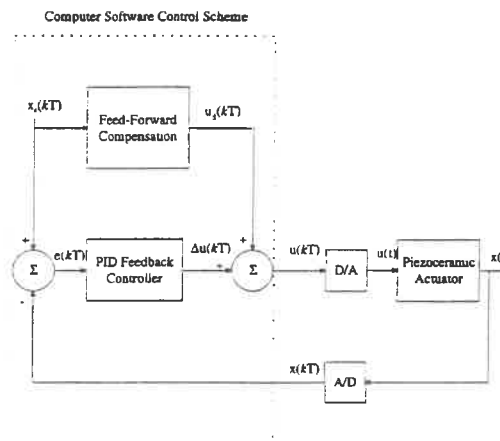


Figure 6. A block diagram of the tracking control system that is made up of a PID feedback controller augmented with a hysteresis linearizing scheme in a feed-forward loop

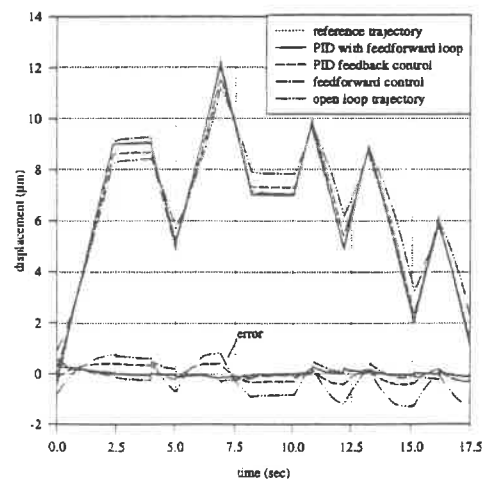


Figure 8. Comparison of four different kinds of tracking control schemes and their errors against the reference trajectory

Implementation of an Acoustic Emission Proximity Detector for Use in Generating Glass Optics*

Mark A. Piscotty, John S. Taylor, Kenneth L. Blaedel

Lawrence Livermore National Laboratory
L-537, P.O. Box 808
Livermore, CA 94551
510-422-2052 fax: 510-423-1460

Background and Motivation

The use acoustic emission (AE) sensing as a method to monitor proximity between a grinding wheel and a brittle material workpiece is being developed at Lawrence Livermore National Laboratory (LLNL) and the Center for Optics Manufacturing (COM) in Rochester, NY¹. Significantly reducing the amount of expensive 'air-grinding' is one of the primary motivations behind this effort, along with lessening the chances of a crash which could damage the wheel, part and machine tool. AE sensing is well developed and routinely used in the metal working industry² for 'initial contact' sensing or tool breakage, for example, and in monitoring diamond turning and grinding processes³. However, using AE sensing to switch from a rapid to a final in-feed rate at the detection of initial-contact between the grinding wheel and a *brittle* material workpiece, such as an optical glass, is often unacceptable during fine grinding (less than 10 μm grit wheels) which produce surfaces with roughness values of 100 Å rms or less. In the approach taken here, we are sensing the AE *prior* to contact between the workpiece and the tool. The coolant between the workpiece and the grinding wheel is used as an AE medium to transfer AE signals generated by the relative motions of the coolant, workpiece and wheel. Capitalizing on the repeatability of the AE approach signal, we have developed a system to detect the proximity of the grinding wheel relative to the workpiece prior to initial contact.

Acoustic Emission System Design at LLNL

We have developed and installed an AE closed-loop feedback system on a spherical optics generator at LLNL. Two air bearing spindles are used on the grinder, one for the grinding wheel and the other for the glass workpiece. A through hole in the part spindle accommodates the AE sensor, AE transmission rod and preamplifier as shown in Figure 1. The transmission rod is a lightly spring-loaded, aluminum rod which is acoustically insulated from the machine tool to minimize machine induced noise in the AE signal. By lightly touching the back side of the workpiece, the transmission rod provides a practical method of coupling the AE sensor to the workpiece, thereby ensuring a desirable signal-to-noise ratio. A rotating union provides the means of connecting the amplified AE signal and preamp power lines between the rotating AE components within the grinding spindle and the outside stationary equipment.

Once the AE signal is transmitted through the rotating union, it is low-pass filtered, passed through an rms converting circuit, into a data acquisition system and finally to the CNC to provide the closed-loop feedback signal. The CNC can then utilize the feedback AE information to perform a variety of commands such as reducing the in-feed rate, changing spindle speeds, etc. Figure 2 is a schematic showing a possible scenario in which two AE thresholds are used to initiate CNC responses. In this scenario, the first threshold would trigger a routine with a set commanded lag time allowing a specified continuation of the fast in-feed (lag time) followed by a ramp down in the in-feed rate. This occurs in the *turbulence region* where the wheel is still not contacting the part and the AE is generated with and

* This work was performed under the auspices of the U. S. Dept. of Energy by LLNL under contract No. W-7405-Eng-48.

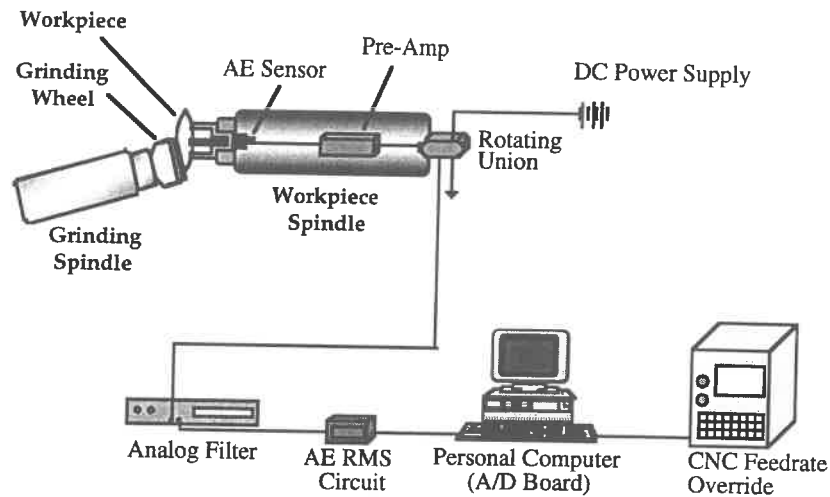


Figure 1. Schematic showing the AE closed loop system at LLNL.

transferred through the coolant layer between the wheel and the part. The next region is referred to as the *once-per-revolution region* in which the wheel is just beginning to make contact with the part due to axial error motions occurring every one revolution of the wheel. This region is noted by a sharp increase in the AE level and the second threshold located in this region could then be used to trigger a quick-change in the feed rate down to the final in-feed rate for the fine grit wheel. And lastly, the final region, *grinding*, is characterized by high, relatively constant AE levels. Repeatability of the AE approach signal is the key which allows exploitation of this process for use as a wheel-to-workpiece proximity sensor.

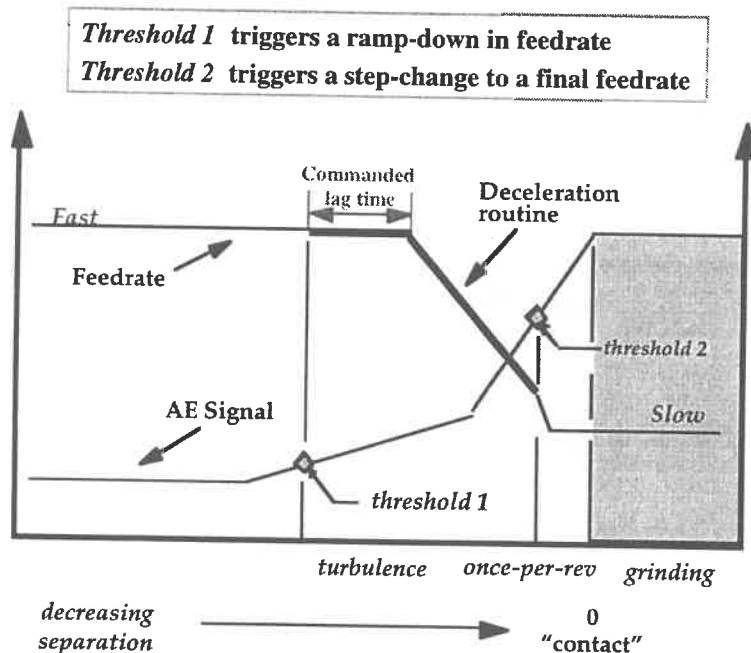


Figure 2. Using two distinct AE level thresholds is a possible implementation method of a closed-loop AE feedback system.

Acoustic Emission Data from LLNL's System

As might be expected with a system like this, it is sensitive to a number of operating parameters such as wheel speed, wheel configuration, coolant flow rate and approach angle of the grinding wheel. For example, Figures 3 a-c, show approach AE data for three grinding situations where the only process parameter change was wheel speed. The data acquisition process began with the approach of the grinding wheel toward the workpiece. Some approach data, such as Figure 3a will have a distinct ridge where the AE signal increases to a certain point and then decreases as the wheel gets closer to contacting the part. Other data will monotonically increase as the wheel-to-workpiece separation decreases, like Figure 3c. From left to right, Figures 3 a-c show approach data for three cases where the wheel speed is decreased. Focusing on only the approach portions of the data, it is evident that the AE levels during the approaches decrease with wheel speed until, in Figure 3c, the signal is almost unusable. The different AE magnitudes of the *contact* points is attributed to thermal growths of the wheel, workpiece and fixturing where a fraction of a micrometer may result in a significant change in the once-per-rev signal, where little material removal takes place. Although these signals may vary as process parameters are changed, when they are held constant, as they typically are during grinding, the AE approach signals are quite repeatable. It is this repeatability of the AE approach signal that allows sensing of the AE signals to be used as a proximity detector. Overcoming different AE approach characteristics as parameters are modified can be accomplished through a quick CNC calibration routine (i.e., recalibrate the system if parameters are changed).

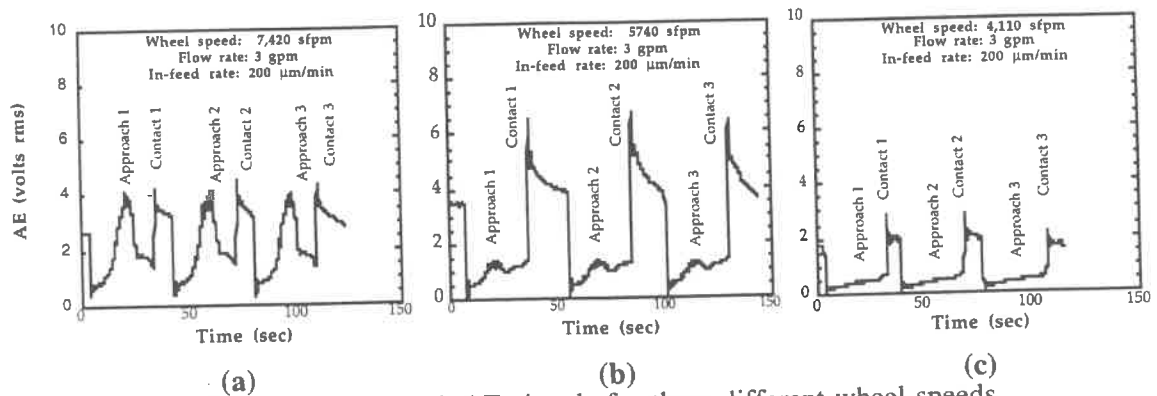


Figure 3. Approach AE signals for three different wheel speeds.

Since the repeatability of the AE system depends on the AE level threshold monitoring, tests were generally conducted using three different in-feed rates, from 50 to 200 μm/min. Figures 4 a and b, show repeatability data for two identical grinding conditions, but with two bronze bond, 2/4 μm diamond wheels with different body configurations. The tests consisted of initially grinding and dwelling on the part to establish the zero point. The wheel is then commanded off the part 100 μm, fed in towards the part until an AE threshold is reached. The CNC then commands the in-feed to stop and the separation distance is recorded.* This process is repeated 20 times to measure repeatability. Adjusting the threshold level will affect the separation stopping distance and can be dialed-in for a desired separation distance.

* Note that this separation is 'implied' from the axis position of the machine tool. Actual separation is sensitive to grinding zone thermal fluctuations which are not reflected in the axis position. This is one source of non-repeatability error in the test data.

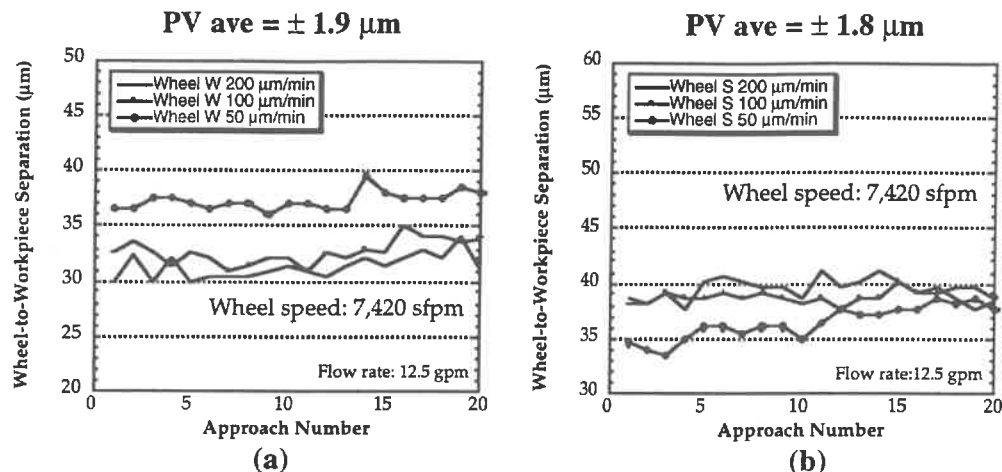


Figure 4. Repeatability data for two different body configurations of wheels.

Figures 5 a and b show repeatability data for two more grinding scenarios. Figure 5a uses the same parameter settings as in Figure 4b, except that the flow rate is reduced from 12.5 gpm to 3 gpm. The result was a more noisy signal which resulted in a greater PV variance. Figure 5b presents two identical repeatability runs, but in this case the wheel speed was 4950 sfpm and the flow rate was 7.5 gpm. This produced a relatively low AE approach level that resulted in a close stopping distance from the part (1 to 2 μm separation). The peak-to-valley variance of this case was by far the lowest of all the cases that corresponded to the approach signal also being the least noisy. The repeatability data for these cases is also a function of sampling strategy (filtering, rms calculation, time averaging) and balancing optimum repeatability and minimum response time is required for each grinding situation.

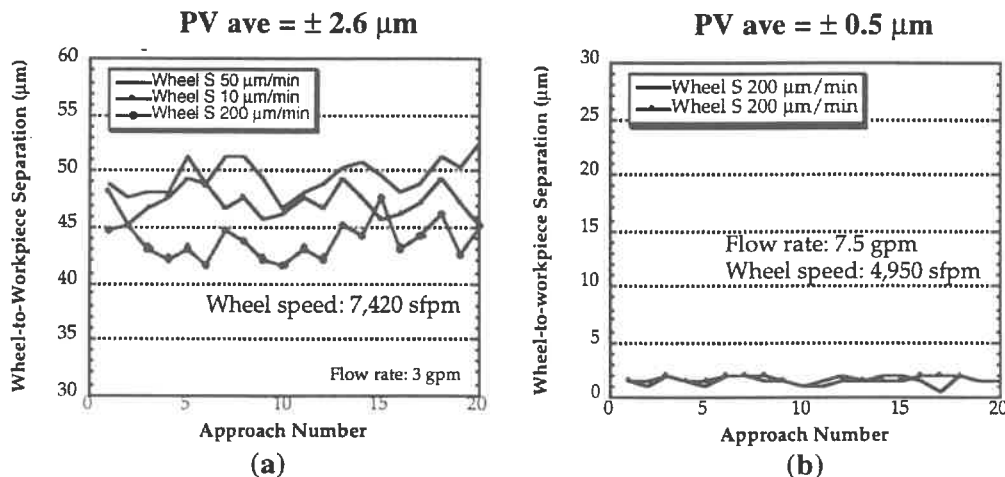


Figure 5. Repeatability data for two different parameter settings with the same wheel.

An additional benefit to the AE proximity system is that it allows real-time monitoring of the grinding process that can reveal such information as the need for the wheel to be dressed. Figure 6 shows an example of this. The upper left plot is of data acquired during a nominal grinding process. After an initial long term dwell, the wheel backs off the part, feeds in, grinds the part and dwells (sparks out) for 45 seconds. As seen in this first plot, the AE signal drops off quickly in the dwell region. The upper right plot is data from an identical run, but this time the wheel is in a loaded condition and needs to be dressed. Data in this plot contains much more noise and the dwell signals remain at relatively high levels. After dressing the wheel, another run is made and the bottom plot shows the resultant data with the return to nominal AE grinding characteristics.

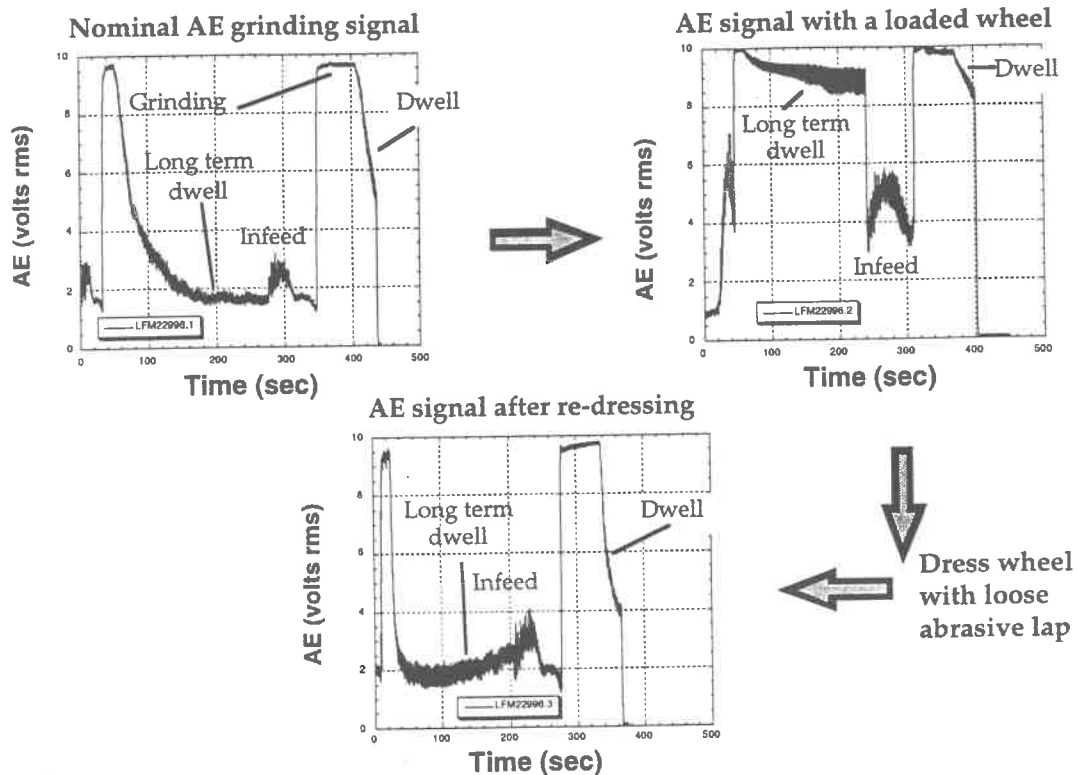


Figure 6. Using AE to monitor in-process grinding data reveals information such as the need to dress a wheel. Bronze bond, 2/4 μ m diamond wheel, grinding BK7 glass.

Commercialization

An AE system similar to the one at LLNL is currently in the development and testing stages on commercially available spherical generator, the Micro SX from CNC Systems⁴, and is located at the COM. The Micro SX has two vertically mounted spindles and can generate parts up to 25 mm in diameter. A significant difference between the LLNL generator and the Micro SX is that the Micro SX utilizes roller bearing spindles which tend to introduce more AE noise into the system. Sufficiently isolating the transmission rod from the workpiece spindle and filtering of the AE signals provides a useable AE signal.

Figure 8 is a close up view of the workpiece chucking mechanism, the transmission rod assembly and the AE sensor. The AE signal line exits the bottom end of the spindle and is connected to the preamp, which is also rotating with the spindle. The AE sensor is mounted on the lower end of the transmission rod and the sensor cable is fed through a center hole in the belt-driven spindle. The amplified AE signal is routed through a rotary union and is brought out to a data acquisition system and the Micro SX's CNC. Special programming of the CNC's programmable logic controller will provide the AE level threshold monitoring necessary for the closed-loop feedback implementation. Initial tests will be used to measure repeatability and robustness of the AE system as a wheel-to-workpiece proximity detector. An additional feature is that real-time in-process information will be available by monitoring the AE signal during grinding.

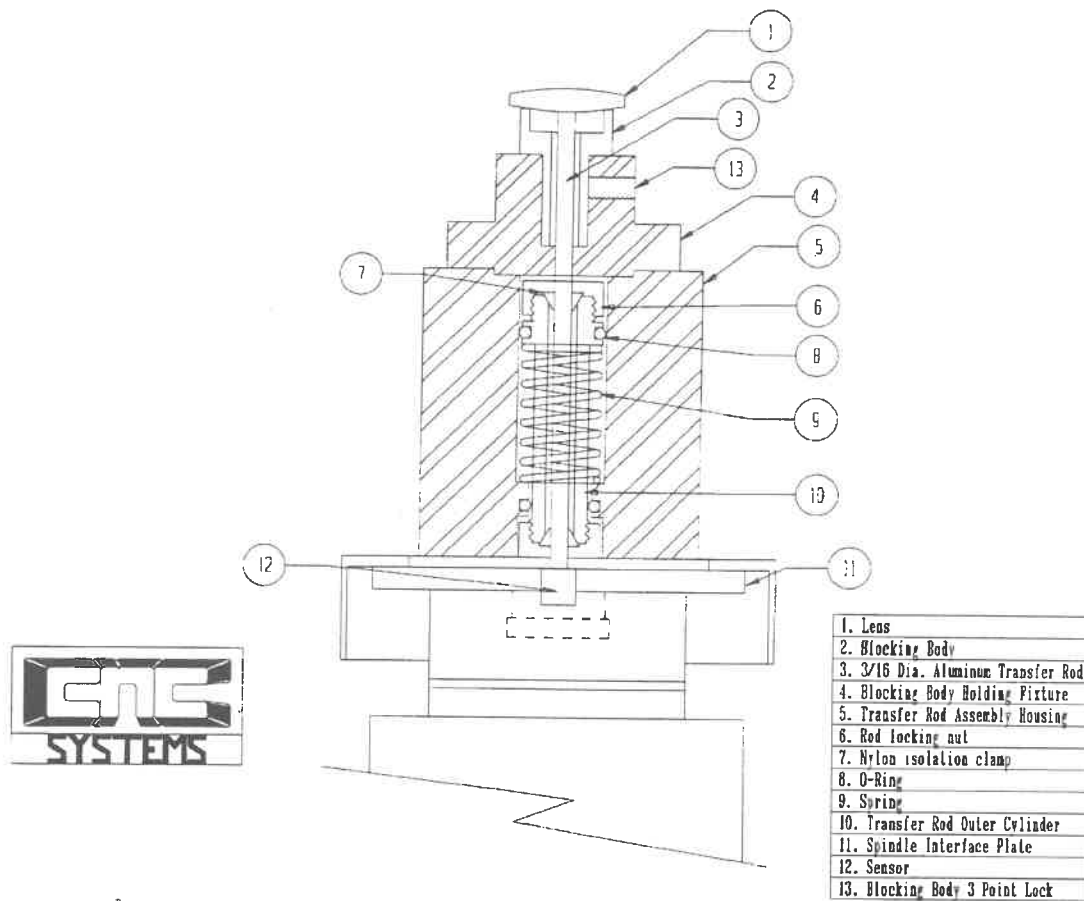


Figure 8. Detail of the acoustic emission sensor fixture for the OptiPro 100.

Summary and Conclusions

We are using the approach AE signal during a grinding operation to detect the proximity of the grinding wheel relative to a brittle material workpiece and are using this detection as a feed-back control signal in our CNC. The repeatability of the AE signal during the wheel approach is the key that allows AE to be used as a proximity detector and is demonstrated at LLNL to be about $\pm 2 \mu\text{m}$. We noted significant changes of the AE signal as process parameters are modified, but conclude that with a quick CNC calibration routine and holding the parameters constant during a given operation, the AE system can be successfully used to sense pre-contact wheel-to-workpiece separation. Additionally, the AE sensing system allows real-time monitoring during grinding to provide in-process information. The first prototype of an AE system on a commercially available generator is currently being tested at the Center for Optics Manufacturing.

¹ Taylor, J.T., Piscotty, M.A., Blaedel, K.L., Dornfeld, D.A., Weaver, L., "Investigation of Acoustic Emission for Use as a Wheel-to-Workpiece Proximity Sensor in Fixed Abrasive Grinding," ASPE Annual Meeting Proceedings, Volume 12, October, 1995.

² Dornfeld, D.A. "Acoustic Emission and metalworking - Survey of Potential and Examples of Applications," 8th NAMRC Conf., May 18-21, 1980, pp 207-213.

³ Bifano, T.G. and Yi, Y. "Acoustic Emission as an Indicator of Material Removal Regime in Glass Micromachining," Journal of Precision Engineering, Vol. 14, No. 4, October 1992, pp 219-228

⁴ CNC Systems, Inc., 369 Route 104, Ontario, NY 14519. Mike Bechtold, President.

Adaptive Pole-Zero Cancellation in Grinding Force Control

Hodge E. Jenkins¹

Georgia Institute of Technology
The George W. Woodruff School
of Mechanical Engineering
Atlanta, GA 30332-0405

Thomas R. Kurfess

Georgia Institute of Technology
The George W. Woodruff School
of Mechanical Engineering
Atlanta, GA 30332-0405

Abstract

Previous work has demonstrated that grinding is a time-varying process, best modeled in real-time. Also it has been shown the integration of multiple sensor data can be used to achieve a superior estimation of a two-parameter grinding model, suitable for adaptive control of the normal grinding force. Building on these developments, this research investigates an adaptive force controller, using the real-time process model with multiple input estimation algorithms as a basis. An adaptive pole-zero cancellation technique is developed to overcome the grinding process variation. With the grinding model estimated in real-time, the adaptive controller provides inverse grinding process dynamics. This cancels the process effects on the controlled system. Real-time estimation and controller designs are implemented using parallel processors for higher speed control capability than is presently available in machine tool controllers. This approach to grinding force control has demonstrated the potential for increasing productivity with the reduction of tary time in plunge grinding, showing higher force regulation fidelity. Decreases in the normal force range and standard deviation is seen in comparisons to fixed-gain controllers. The implementation of this adaptive pole-zero cancellation approach in surface (traverse) grinding provides a superior surface following capability for fine finishing, as compared to fixed gain controllers.

Background

Grinding is a complex operation, seen as three separate and concurrent process actions: (1) cutting, (2) plowing, and (3) rubbing (Hahn and Lindsay 1971), occurring simultaneously in varying distributions depending on the grinding wheel, material, and operating conditions. This causes detailed process analysis, modeling and control to be difficult. Previous work has demonstrated that grinding is a time-varying process and should be modeled in real-time. The empirical grinding process model selected in this research is based on a combination of the grinding model developed by Hahn and Lindsay as well as the Preston model (Jenkins, et al, 1996). The selected, two-parameter grinding model is given in equation (1).

¹ Author is presently with Lucent Technologies, Bell Laboratories in Norcross, GA.

$$Q = K_p(F_N - F_{TH})V; \quad F_N \geq F_{TH} \quad (1)$$

The material removal rate is defined as Q , while F_N is the applied normal force, and F_{TH} is the threshold normal force. V is the surface speed of the grinding wheel. The coefficient, K_p , and the threshold force, F_{TH} , are determined experimentally.

Although fixed-gain controllers (such as PID-type controllers) are preferred for industrial applications because of their simplicity in implementation and tuning. However, many processes or systems vary sufficiently to limit the usefulness or stability of fixed-gain controllers. Grinding has been shown to be a time varying process that can benefit from the application of such a controller.

The desired objective of this research is to provide a methodology for a stable force control in grinding to reduce cycle time and maintaining geometry in fine finishing, despite grinding process variation. The adaptive controller uses pole-zero cancellation and direct design methods to be self-tuning. The design approach and stability evaluation of the adaptive force controller, based on the previously described real-time process model, are presented.

Adaptive Process Control

To better control the normal grinding force for the system, adaptive control methods are applied building on the successful real-time grinding model with on-line parameter estimation. The approach taken in this adaptive force control effort for grinding is different from that of Tonshoff et al (1986). Tonshoff applied a fixed-gain controller to regulate the grinding force for an internal diameter grinder. Tonshoff used a separate off-line model estimation routine to determine the fixed-gains of the controller. In contrast the work undertaken here performs real-time estimation of the grinding process and uses a pole-zero cancellation approach to remove the process variation.

A block diagram of the overall system is depicted in Figure 1. The system model consists of several functional blocks: force controller, $G_{FC}(s)$, the position control loop, $G_{PC}(s)$, the tool-workpiece process interaction, $G_{TWP}(s)$, as well as the estimation and adaptation functions.

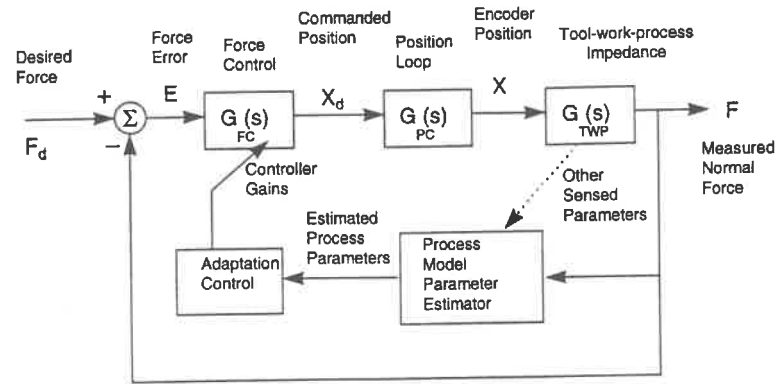


Figure 1. Plant and Controller Schematic for PMAC-based Force Control.

The development of the adaptive force controller for the grinding process can proceed directly from the plant model of Jenkins (1996).

$$G_{TPW}(s) = \frac{F_N(s)}{x_e(s)} = K_s \left(\frac{s}{s + \eta} \right) \quad (2)$$

where $\eta = K_p K_s V/A$

This is combined with a first order position servo model to yield the following continuous plant transfer function.

$$G_{PC}(s)G_{TWP}(s) = \frac{F_N(s)}{x_d(s)} = 203K_s \left(\frac{s}{s + \eta} \right) \frac{1}{(s + 224)} \quad (3)$$

It has been shown that $\eta \ll 224$. Thus, variations in the process can dominate the system response. A fixed-gain controller can lead to poor performance as well as instabilities. An adaptive controller is desired to mitigate the effects of changing process dynamics and maintain a high MRR and a stable controller. From equation (3) it can be seen that the process dynamics can be eliminated by a pole-zero cancellation using a PI-type controller. The controlled system would then be of the form

$$G_{FC}(s)G_{PC}(s)G_{TWP}(s) = \frac{F_N(s)}{E_F(s)} = 203K_s \left(\frac{s}{s + \eta} \right) \frac{K_{PROP}(s + \alpha)}{s(s + 224)} \quad (4)$$

In this scheme if the α (or K_I/K_{PROP}) of the PI controller in equation (4) is set to η , then the entire process grinding dynamics can be eliminated from the resulting equation. This reduces equation (4) to that of only a contact stiffness with a position loop (5).

$$G_{FC}(s)G_{PC}(s)G_{TWP}(s) = \frac{F_N(s)}{E_F(s)} = \frac{203K_s K_{PROP}}{(s + 224)} \quad (5)$$

where $\alpha = \eta$. Since a discrete equivalent system is needed for the sampled plant in order to implement such a scheme on the digital controller, a ZOH equivalent discrete system is used.

$$G_{PC}G_{TWP}(z) = \frac{F_N(z)}{X_d(z)} = \left(\frac{z-1}{z} \right) Z \left(\mathcal{L}^{-1} \left[\frac{1}{s} 203K_s \left(\frac{s}{s+\eta} \right) \frac{1}{(s+224)} \right] \right) \quad (6)$$

denoting

$$C_o = e^{-224T}; f(\eta) = e^{-\eta T}$$

$$f_2(\eta) = \frac{203K_s}{224 - \eta} (f_1(\eta) - C_o) \quad (7)$$

$$G_{PC}G_{TWP}(z) = f_2(\eta) \left(\frac{z-1}{(z-f_1(\eta))(z-C_o)} \right) \quad (8)$$

The discrete PI-type controller can be rewritten as follows

$$G_{FC}(z) = \frac{K_{PROP}}{1+\alpha} \left(\frac{z - \frac{1}{(1+\alpha)}}{z-1} \right) \quad (9)$$

Thus if η varies sufficiently slowly or the controller is sufficiently fast, α can be chosen such that

$$\alpha = \frac{1}{f_1(\eta)} - 1 \quad (10)$$

then complete open-loop discrete transfer function is defined as

$$G_{FC}G_{PC}G_{TWP}(z) = \frac{K_{PROP}}{1+\alpha} f_2(\eta) \left(\frac{1}{z-C_o} \right) \quad (11)$$

Equations (7), (10), and (11) are used to determine the PI force controller gains in the controller design update of the self-tuning regulator. The effectiveness of the pole-zero cancellation is related to the process variation frequency spectrum and the gain update rate.

The implementation of this controller uses the windowed data handling scheme discussed by Jenkins (1996). The size of the data window is the key parameter to the success of this approach. The smaller the window the faster the gain update. However, as the sample window gets smaller the variance of the parameter estimation increases. This reduces the confidence of the estimate and ultimately the reduces the probability of an effective pole-zero cancellation.

A rigorous stability analysis in the sense of Lyapunov is complicated for STR systems, as the controllers are highly nonlinear. The relationship of the regulated force to the estimated parameter can be quite convoluted and is difficult to attain in many instances. A more tangible evaluation of stability can be performed, by a parametric examination of the effect of poor estimation and adaptive performance. The effect of poor performance can be seen by examining the combined open-loop, discrete plant and controller transfer functions given in equations (9) and (11).

$$G_{FC}G_{PC}G_{TWP}(z) = \frac{K_{PROP}f_2(\eta)}{1 + \alpha} \left(\frac{z - \frac{1}{1 + \alpha}}{(z - f_1(\eta))(z - C_o)} \right) \quad (13)$$

From equations (9) and (10) it can be seen that the poles of the open-loop system are all within the unit circle, so they are stable. The variable, α , is the ratio of the integral to the proportional gain and is always positive. Therefore the transfer function zero (4) is always positive and less than or equal to one and is at worst marginally stable.

$$z = \frac{1}{1 + \alpha} \quad (4)$$

Another possibility for instability is in the calculation of the proportional gain. The overall gain is positive and is calculated in real-time to be stable, dependent on the estimated grinding model. However, α is positive and determined by the controller. Thus even if the pole-zero cancellation is not accurate the effects do not cause the system to go unstable; however, performance may suffer.

Results

Two benchmark experiments were conducted to evaluate the effectiveness of the adaptive pole-zero cancellation control and RLS-FF estimation technique. In the first experiment, plunge grinding, the adaptive force regulation successfully reduces the standard deviation and range of the controlled normal force, and yields a mean normal force closer to the 25 N reference. The second

benchmark experiment, a traverse grind over a flat surface, demonstrates better repeatability of the surface profile after grinding. The adaptive control and estimation technique appears to provide an improvement in surface following for fine finishing, compared to fixed-gain controllers. The ratio of several surface profile characteristics is provided in Table 1.

Table 1. Traverse Grind Surface Profile Statistic Ratios.

Material	O1	A2	4142
<u>Ratio of Adaptive Control to Fixed-Gain Control</u>			
Absolute Mean	35.7%	37.3%	70.2%
RMS mean	36.4%	40.9%	66.6%
Range	47.1%	54.7%	68.3%
Std. Deviation	36.4%	40.9%	66.6%

Conclusion Summary

A self-tuning regulator approach to adaptive control has been successfully developed for force control of a grinding system. The controller is based on a pole-zero cancellation of the estimated grinding process. The use of parallel processors is a key element in the combined servo control and process estimation functions. The final implementation of the adaptive controller and estimator yielded a superior grinding force controller. A higher fidelity force regulator with greater stability has been developed for plunge grinding. Similarly in traverse grinding, the surface profile following capability improves with the adaptive pole-zero cancellation technique, as contrasted to the fixed-gain controller. This approach to grinding force control has demonstrated the potential for increasing productivity with the reduction of tary time by use of the stable force control in plunge grinding, with higher fidelity in force regulation. The implementation of this approach in surface (traverse) grinding provides a superior surface following capability for fine finishing, as compared to fixed gain controllers. These conclusions are evident from the results of the benchmark experiments.

BIBLIOGRAPHY

- Hahn, R. S. and Lindsay, R. P. "Principles of Grinding Part I Basic Relationships in Precision Machining," *Machinery*, July 1971, pp. 55-62.
- Jenkins, H.E., Kurfess, T.R., and Ludwick, S.J., "Determination of a Dynamic Grinding Model," (accepted February 1996) *ASME Journal of Dynamic Systems, Measurement and Control*.
- Landau, I. D., *Adaptive Control*, Dekker, New York, 1979.
- Tonshoff, H. K., Zinngrebe, M., Kemmerling, M., "Optimization of Internal Grinding by Microcomputer-Based Force Control," *Annals of the CIRP*, Vol. 35, No. 1, 1986. pp. 293-296.

DIAMOND POINT TOOL FORCE MEASUREMENT IN A PRODUCTION ENVIRONMENT

J. Schaefer (Texas Instruments, Inc.), **C. Arcona** (IBM Corporation), **T.A. Dow**, **G.M. Moorefield II** (North Carolina State University), **K. Falter** (Rank Pneumo, Inc.), and **R.L. Rhorer** (National Institute of Standards and Technology)

INTRODUCTION

Diamond turning is used to produce optical surfaces in ductile and brittle workpieces. Materials of interest include aluminum, copper and electroless nickel for reflective optics and polycrystalline Germanium for transmissive infrared lenses. The geometry of the diamond tool plays an important role in the resulting surface finish and figure error for these optics. However, it is difficult to predict the change in geometry due to wear, the effect of wear on the surface and, most important, the appropriate time to change tools to optimize the cost and quality of the optics.

Tool force can be used to monitor the quality of the machined surface. Changes in force can forecast a change in surface finish due to a shift in the minimum chip thickness as well as a change in figure accuracy as the machine elastically deforms in response to increased tool forces. Although measuring tool forces during production has been considered a desirable

goal, it has remained unimplemented because of (1) compliance typically introduced by including the force sensor in the structural loop and (2) the lack of availability of a user-friendly interface to assist the operator in making process and tool change decisions. To address this problem, a collaborative project has been undertaken by Los Alamos National Labs (LANL), Texas Instruments (TI) and North Carolina State University (NCSSU). The objective of this project is to develop a tool force dynamometer and collect cutting force data in the background of normal part machining and display information about the process, tool condition, and tool life to the operator.

PRODUCTION FORCE ANALYSIS TOOL

The Production Force Analysis Tool (PFAT) was designed around a Kistler 9251A three-axis load cell which uses three pairs of quartz transducer rings to resolve forces in three orthogonal directions via the piezoelectric effect

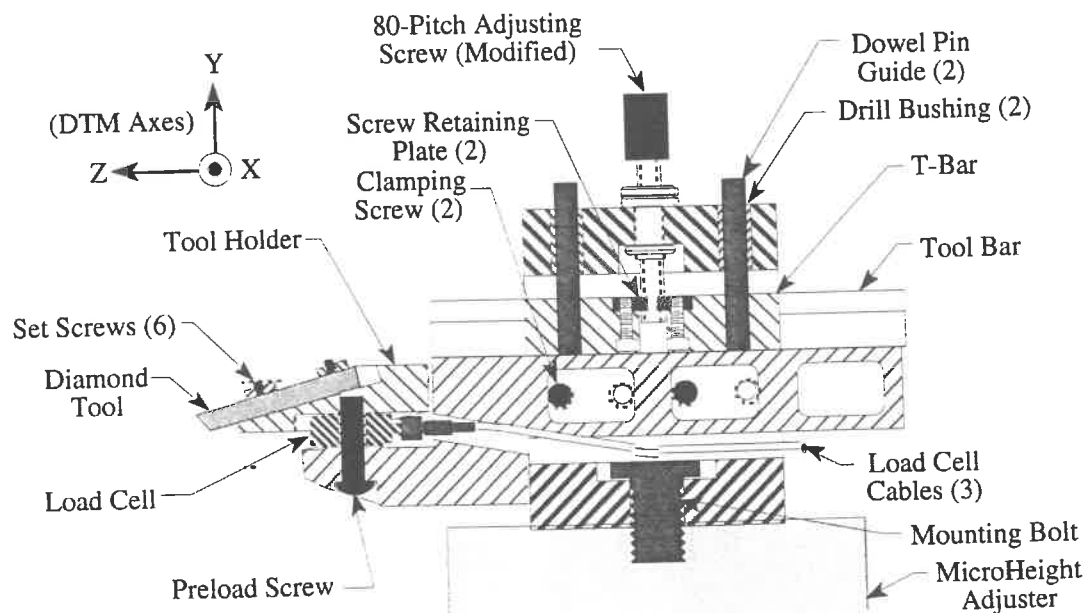


Figure 1: Cross-sectional view of the PFAT prototype.

(i.e., stresses in the quartz generate proportional charges). For reliable force measurement, the dynamometer must provide a rigid platform to which the load cell can be mounted, and a compact, yet robust, tool holder that can be attached to the load cell with the necessary preload. For this low-force application, a preload of approximately 8900 N was chosen. Because this preload force generates a 0.4 GPa tensile stress in the screw that clamps the load cell between the tool holder and the large cross-section tool bar, a heat treated alloy steel screw ($\sigma_{ut} \geq 1.17$ GPa) was selected.

To accommodate a wide range of workpiece thicknesses, this dynamometer design provides 40 mm of axial tool motion. The axial tool position is adjusted by first loosening the two clamping screws that constrain the moveable tool bar between the sides of the mounting structure, manually sliding the tool bar forward or backward, and then retightening the clamping screws to secure the tool bar in position. The dynamometer design also provides 12.7 mm of vertical tool positioning to allow coarse tool height alignment when tool changes are made. To adjust the tool height (see Figure 1), the two clamping screws are loosened to free the tool bar and the 80 pitch adjusting screw is used to vertically position the T-bar. To provide fine vertical tool positioning, the dynamometer is mounted on a MicroHeight Adjuster. The incorporation of the tool positioning mechanism in the dynamometer design eliminates the need for an Aloris tool post which is a weak link in the structural loop of the machine.

Tests were conducted to determine the static stiffnesses and the first resonant frequencies in the axial, horizontal, and vertical directions of a tool mounted in the PFAT. Supporting the tool with a stiff structure is necessary for the

dynamometer to produce components with good figure and finish. High first resonances of the structure enable the load cell to operate with greater bandwidth, permitting the observation of high frequency force fluctuations. Table 1, which compares the experimental results with the design values, reveals excellent agreement between modeled and measured stiffness in all three directions.

The fact that all of the first resonant frequencies are in excess of 1000 Hz suggests that the cutting forces can be successfully monitored and correlated with the condition of the tool edge (i.e., the amount of tool wear). Furthermore, when FE models of the PFAT prototype and TI's production-proven tool holder are compared, the results indicate that the two structures possess nearly identical first resonant frequencies, despite the fact that the PFAT prototype includes the load cell while the TI device does not. The FE models indicate that the two structures have the same first vertical resonance (1700-1800 Hz) and the same first horizontal resonance (1800-1900 Hz). The experimental resonances of the PFAT prototype are lower than the predicted frequencies because the FE models did not include the tool post and microheight adjuster which reduce the natural frequencies.

PRODUCTION ENVIRONMENT

For the PFAT design to be beneficial in a production environment, the system must provide the following: (1) similar adjustments as the existing production tool holder, (2) static compliance that meets or exceeds that of the existing tool holder, (3) a data collection system that requires minimal operator interaction, (4) the data analysis software must be robust for accurate force data extrapolation, and (5) software that displays both real time, and historical tool force data to the operator for

Direction	Stiffness (N/mm x 10 ⁻³)		Natural Frequency (Hz)	
	FEM Model	Measured	FEM Model	Measured
Axial (Z)	27	28	-	2275
Horizontal (X)	3.7	3.3	1850	1275
Vertical (Y)	2.8	2.7	1750	1075

Table 1. Predicted and measured stiffness and resonance of the PFAT prototype

tool condition and process assessment.

There are multiple issues that naturally occur in a production environment that must be addressed in order for the tool force to be consistent, credible, and useful. Some of the issues that were considered when designing the tool force system and software included: (1) low run size, high part type mix, (2) multiple shift operation and operational variations, (3) influence of spray-mist coolant, (4) curvature of lenses, (5) process variation, and (6) different workpiece materials.

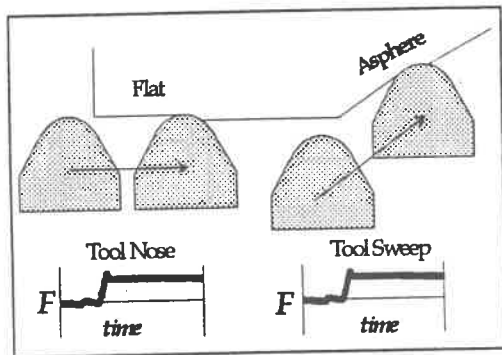


Figure 2. Tool edge condition assessment

Tool Condition Assessment Scheme

A fundamental limitation of a piezoelectric force transducer is that the force signal dissipates over time thus preventing continuous force monitoring of the entire finish pass. Most of the aspheric lenses machined at TI are concave with an outer locating flat. For this geometry, a method was devised to monitor the conditions of the nose and the outermost regions of the tool sweep for each finish pass. Tool force data recorded on the flat surface as shown in Figure 2 will register changes at the center of the tool while forces recorded as the tool enters the edge of the asphere will be influenced by tool changes at the other end of the tool sweep.

Data Collection and Software

To collect tool force data, a National Instruments data acquisition card was installed in a 486PC. The key to successful implementation of production tool force measurement lies in the software, and for this application, LabWindows CVI version 3.1, C

language software was used. It was imperative that the data acquisition and software systems perform the work of a lab technician by extrapolating, validating, and presenting tool force information to the operator in an understandable format.

Because the tool forces are so small, it is a challenge for the software to detect and identify the force transition from an unloaded to a fully loaded tool. One source of confusion is the noise created when a the spraymist coolant is used. Several different software schemes have been evaluated to identify this transition zone including the change in the average force and the change in standard deviation of the thrust force from no cutting to cutting.

Data Analysis Figure 3 is an example of a tool force profile with coolant on a polycrystalline Ge workpiece using a 0.81 mm nose radii tool with a 1.25 μm depth of cut and a 1.25 μm feed per revolution. The depth of cut variation from initial tool contact to full depth causes the gradual transition slope. The data is averaged over a half-second interval before and after the transition to obtain the average cutting and thrust forces. A change in the resultant force angle between F_{cutting} and F_{thrust} indicates tool edge wear. As the tool wears, F_{thrust} increases while F_{cutting} remains nearly constant, thereby increasing the angle of the resultant force $\alpha_{\text{resultant}}$, where

$$\alpha_{\text{resultant}} = \tan^{-1} \left(\frac{F_{\text{thrust}}}{F_{\text{cutting}}} \right) \quad (1)$$

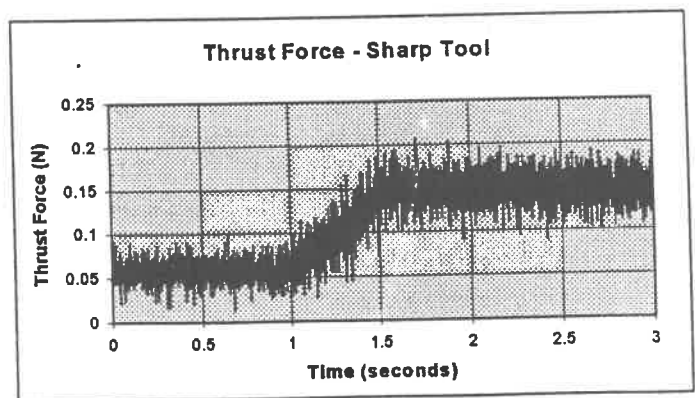


Figure 3. Thrust force data for Germanium

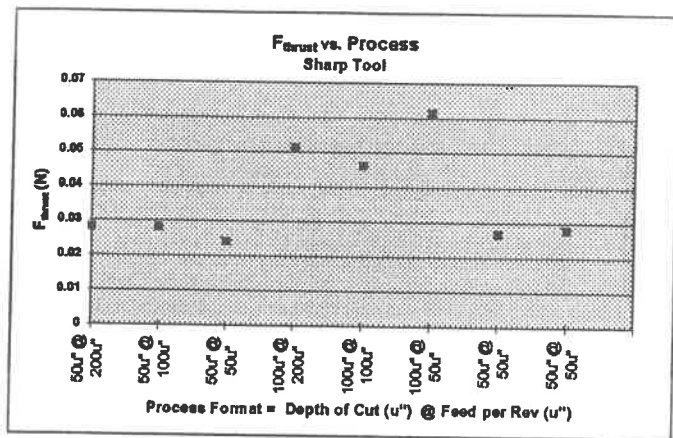


Figure 4. Thrust force for different conditions

Change in the standard deviation of the force signal from no cutting to cutting might also provide important information on the process such as the amount of material fracturing during machining. The chip area calculation is used to normalize cutting forces to account for process changes which affect the measured cutting forces and the chip volume calculation is used to track tool wear rates and tool change out thresholds.

RESULTS

The standard machining process for Germanium can be characterized using the measured tool force and chip structure. The typical process utilizes a negative 15 degree rake angle tool with a 0.76 mm nose radius, 1.25 μm depth of cut with a 1.25 μm feed per revolution produced by a feed rate of 1.9 mm per minute at a spindle speed of 1500 rpm. From Figure 4, it is clear that feed per revolution has little effect on the thrust force, whereas, doubling the depth of cut nearly doubles the thrust force. This is due to the proportional increase in the contact area of the tool as the depth of cut increases.

SEM micrographs of the Germanium chips confirm the ductile nature of this process. The tool used in this particular experiment had removed 75 μm of material from 12, 76.2 mm diameter Germanium lenses. The measured forces were $F_{cutting} = 0.069$ N, $F_{thrust} = 0.024$ N. In Figure 5, the thickest portion of the chip is at the bottom of the picture and the chip is thinnest where it appears to form strands.

These strands are actually hollow tubes similar to a rolled up piece of paper. The theoretical cross-sectional chip geometry is 0.075 μm thick by 40 μm wide. This correlates well with the scaled width in the micrograph.

SUMMARY AND DISCUSSION

Production diamond tool forces (~ 0.05 N) are smaller than conventional machining forces. This creates difficulty in extracting accurate cutting force data when small external disturbances generate variations in the measured values. For example, the standard spraymist application of coolant creates noise in the tool force data.

Filtering the data and/or changing to a steady stream of coolant will be used to eliminate this difficulty. The thrust force on a sharp tool is nearly proportional to the depth of cut and is affected very little by a change in feed. Furthermore, changes in the standard deviation, σ , of the force signal is greater with a process creating fracture than a ductile process. The σ parameter could prove to be invaluable in process optimization. Another benefit of on-machine tool force measurement is the increased accuracy (< 20 nm) afforded the operator in registering the tool with the part surface. In conclusion, by providing proper hardware, software, and training to the operator, tool force measurement in a production diamond turning operation is viable and has proven to be an invaluable tool.

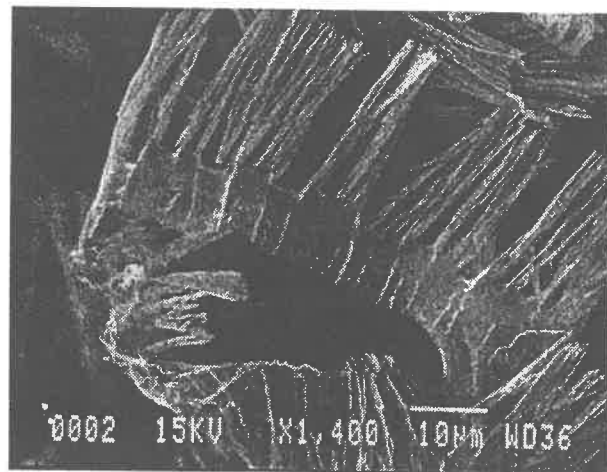


Figure 5. Germanium chip

Submicrometer Overshoot Control of Rapid and Precise Positioning

S. B. Chang

Precision Instrument Development Center, National Science Council

Hsinchu, Taiwan, R.O.C.

I. Introduction

Rapid and precise positioning is an important technique in micromachining such as in the machining and processing of semiconductors [1]~[3]. Also, small overshoot or submicrometer overshoot requirement is a main concern of controller design in many micromachining cases such as micro cutting, milling and turning. Therefore, to achieve submicrometer overshoot in a rapid and precise positioning is an important technique. The control algorithm proposed in this paper adopts the variable structure control method to provide a rapid, precise and submicrometer overshoot positioning. It is implemented on a linear motor drive stage guided by a rolling ball guide and with a 10 nm resolution laser interferometer as the displacement sensor to achieve an overshoot smaller than 0.2 μm and rapid, precise positioning (1mm ~ 1cm positioning with steady state error smaller than 30 nanometer in 0.5 sec). The controllers in this algorithm are designed according to the corresponding dynamics and frictional effects of the stage during a long stroke positioning motion. It shows by combining mechanical design, system identification, lubrication and controller design techniques will lead to more successful results.

II. System Identification

The functional block diagram of our positioning system is shown in Figure 1. Similar to the experimental results in [2], we found the open loop frequency response of the plant varies with the magnitude of the input voltage. The dynamics and the corresponding transfer functions of the plant identified with different input voltages are summarized in Table 1.

III. Dynamics Transition During a Motion

In order to decide the dynamic condition of the plant during a motion, the following two definitions are proposed.

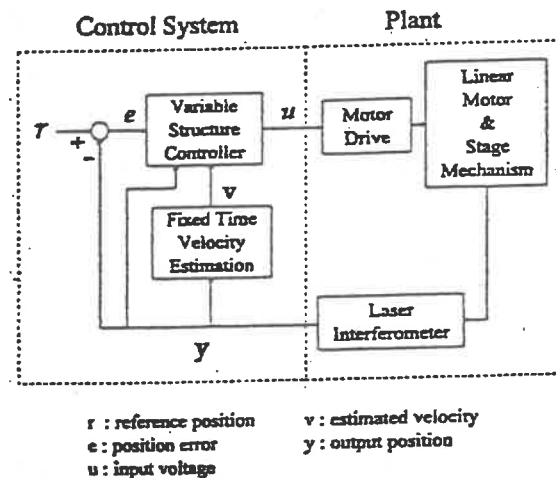


Figure 1. Functional block diagram of our positioning system.

Input voltage	0.05 V	0.2 V	0.5 V
dynamics	a hard spring element	a simple mass-damper-spring system	an ordinary 2nd order Type 1 system
corresponding transfer function	$\frac{ss}{s^2 + 100s + 6000}$	$\frac{ss}{s^2 + 12.5s + 700}$	$\frac{ss}{s(s+1)}$

Table 1. Dynamics and corresponding transfer functions of the plant for different input voltages.

Definitions:

a. Stick Point :

When the input force of a stage is removed, the point where the stage naturally stops and sticks is called a 'stick point'.

b. Critical Speed :

We define the speed below which the stage without input force will naturally stops and sticks within the distance of the maximum presliding displacement. According to the conservation of energy and neglecting the dissipative property of friction, we obtain the critical speed as follows

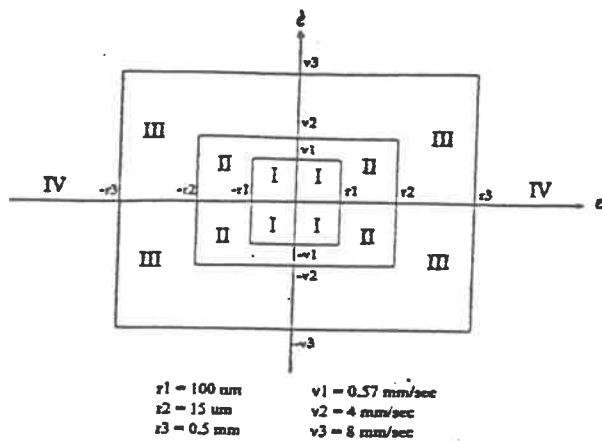


Figure 2. Regions division of the proposed variable structure control algorithm.

$$V_{cr} = \sqrt{\frac{2 \int_0^{pd} f(x) dx}{M}} \quad \dots (1)$$

where pd is the maximum presliding displacement, f is the friction under stick condition, M is the mass of the stage and x is the displacement.

The critical speed of the stage in our experiments is about 0.57 mm/sec. Then a dynamics transition condition is obtained as follows.

Dynamics Transition Condition

When the speed of a stage varies from a value larger than the critical speed to a value smaller than the critical speed, the dynamics of the stage varies from macrodynamics to microdynamics and a stick point occurs at that instant. In contrast, when the speed of a stage varies from a value lower than the critical speed to a value larger than the critical speed, the dynamics of the stage varies from microdynamics to macrodynamics and a stick point disappears at that instant.

In the precise positionings of long stroke, the stage must go through three different kinds of dynamics as described above. Moreover, the final precise positioning is fulfilled by using the presliding displacement of a stick point. Hence, in order to achieve a submicrometer overshoot, rapid and precise positioning, the following variable structure control algorithm is proposed.

IV. Proposed Control Algorithm

The control algorithm proposed in this paper divides the positioning range into 4 regions, as shown in Fig.2, according to the state of the stage.

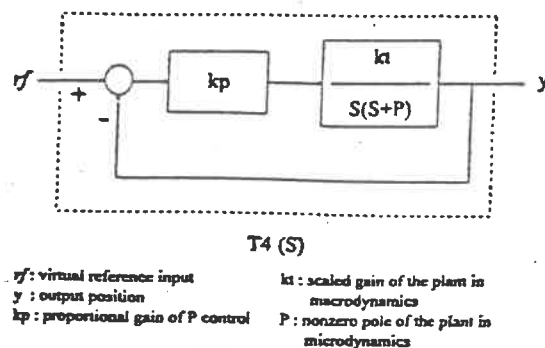


Figure 3. Block diagram of the virtual reference control in Region IV.

When the stage far from the reference position (position error > 0.5 mm or the moving speed > 8 mm/sec, Region IV), a novel and simple control method we called 'the virtual reference control' is proposed to let the stage move rapidly toward the destination position and decreases the speed near zero when it moves into the region III (position error ≤ 0.5 mm and speed ≤ 8 mm/sec). Then a constant speed control is adopted, which is proved that stick-slip will not occur and it can let the stage move into region II (position error ≤ 15 μm and speed ≤ 4 mm/sec). In region II, pole assignment control with small integrator gain is adopted to achieve submicrometer overshoot. And finally, when the stage moves into region I (position error ≤ 100 μm and speed ≤ 0.57 mm/sec), pole assignment control with large integrator gain is adopted to shorten the settling time.

a. Virtual Reference Control in Region IV

The control configuration in Region IV is shown in Figure.3. If the destination position is denoted as r , we set the reference input rf to $T4(S)$ as follows:

$$rf = \frac{r - \text{sign}(r) * r1}{1 + u * Mp} \quad \dots (2)$$

where $r1$ is the width of Region I, u is the adjusting factor and Mp is the maximum overshoot of the system to a step response [4], $\text{sign}(r) = 1$ if $r > 0$, $\text{sign}(r) = -1$ if $r < 0$.

By using above rf value as the reference input of $T4(S)$ and the value of u is 1, then we can obtain the top displacement of the response is about $r - \text{sign}(r) * r1$. Therefore, this control algorithm can move the stage rapidly and smoothly into Region III.

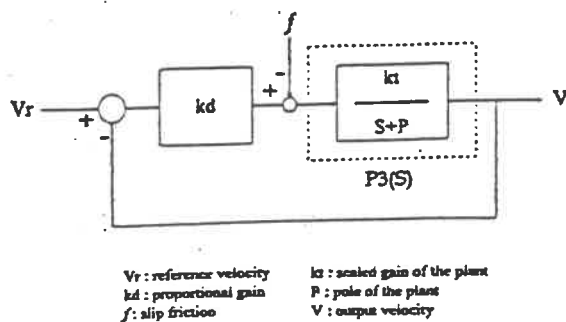


Figure 4. Block diagram of the constant speed control in Region III.

b. Constant speed control in Region III

To simplify the analysis, we regard our speed servo system in Region III as an ordinary speed servo system suffering friction as shown in Figure 4. Where V_r is the reference speed, k_d is the proportional gain and f is the scaled friction whose value is between $f_{\min}$ and $f_{\max}$. If the reference speed V_r is chosen to be large enough to avoid the occurrence of stick-slip, the velocity V presented in the time domain will be limited as follows:

$$V_{\max}(t) \geq V(t) \geq V_{\min}(t) \quad \dots (3)$$

$$V_{\max}(t) = [V_r - k_t * f_{\min} / (P + k_d k_t)] (1 - e^{-(P + k_d k_t)t})$$

$$V_{\min}(t) = [V_r - k_t * f_{\max} / (P + k_d k_t)] (1 - e^{-(P + k_d k_t)t})$$

Therefore, by properly setting the parameters in Region III, the constant speed control can move the stage smoothly toward Region II without the occurrence of stick points near Region II. And it can limit the speed below $V_{\max}$ in Eq.(3) when the stage moving into Region II.

c. Pole assignment control with small integrator gain in Region II

In Region II the pole assignment control is adopted in order to achieve small overshoot requirement. Its control configuration is shown in Fig.5 and the overall transfer function can be expressed as follows,

$$T2(S) = \frac{a b c}{(S + a)(S + b)(S + c)} \quad \dots (4)$$

Then according to the pole assignment control rule [4], we can arbitrarily assign the poles of $T2(S)$, i.e., $-a, -b$ and $-c$, by choosing certain k_i, k_s and k_v values.

Then the time response of the velocity $y'(t)$ in the final positioning can be simplified as follows:

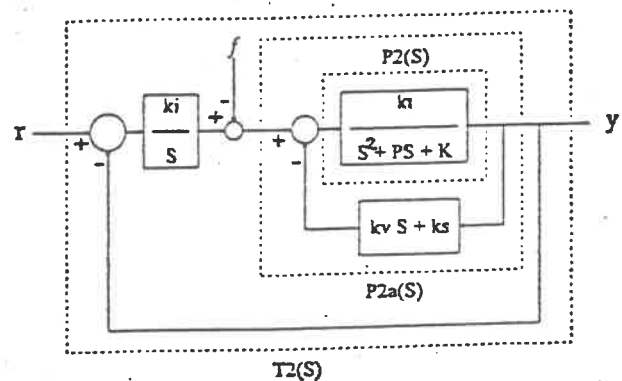


Figure 5. Block diagram of the pole assignment control in Region II and Region I.

$$y'(t) \approx C e^{-a} \quad \dots (5)$$

if positive real parts of $a, b \gg c$

$$C = \frac{1}{(b-c)(c-a)} [c(a+b)y'(0) - a b c r]$$

The term e^{-a} in Eq.(5) varies from 1 to 0 as time goes from 0 to infinity. Hence, to achieve no velocity reversal and no overshoot in the final positioning we have to let C in the Eq.(5) be positive. Moreover, in order to keep the stage in microdynamics and avoid stick-slip in the final positioning, we have to let C in the Eq.(5) smaller than the critical speed V_{cr} . Because the real parts of a and b are larger than c , $(b-c)(c-a)$ is negative. Therefore, to result in no overshoot and no stick-slip motion in Region II, the following equation must be satisfied.

$$\frac{(b-c)(c-a)V_{cr} + a b c r}{c(a+b)} < y'(0) < \frac{a b}{a+b} \cdot r \quad \dots (6)$$

Therefore, if we can properly set the poles locations $-a, -b$ and $-c$ and limit the maximum value of $y'(0)$ by the constant speed control in Region III, a no overshoot positioning can be achieved theoretically.

d. Pole assignment control with large integrator gain in Region I

When stage moves from Region II into Region I, the position is near the destination position and the velocity is below the critical speed in Eq.(1). Therefore, in Region I the stage is in microdynamics and if we choose a properly large integrator gain k_i ,

TORSIONAL COMPLIANCE IN GRINDING CHATTER

Rodney D. Entwistle and Brian J. Stone
The University of Western Australia
Nedlands, Perth 6009
Australia

Introduction

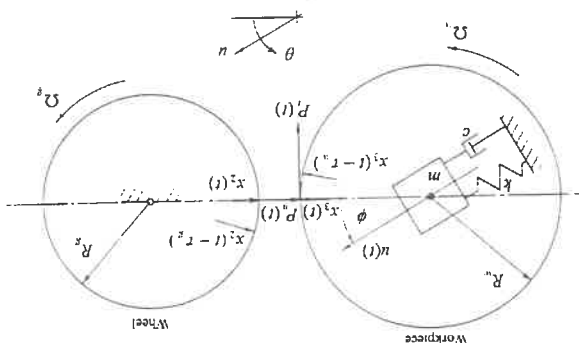
Grinding is one of the important finish machining processes available for the manufacture of components requiring fine tolerances and/or fine surface finishes. However, the grinding machine tool is potentially subject to dynamic instability, known as chatter, which when present can reduce production rates, lead to a deterioration in surface finish, reduce the time between wheel dressings and cause aural discomfort for the machinist.

Research into chatter theory in general and grinding chatter in particular has been extensive, leading to a reasonably good understanding of the parameters involved. In the case of cylindrical grinding both the workpiece and the grinding wheel are rotating systems which will exhibit torsional responses. It appears that the grinding chatter models and simulations developed to date have assumed either explicitly or implicitly that torsional vibrations will not have any affect on the stability of the grinding process. This may seem reasonable when it is considered that torsional vibration of the workpiece will have an extremely small affect on the chip thickness. However, the cutting force models generally in use indicate that the tangential contact force in grinding is proportional to the workpiece surface speed and inversely proportional to the grinding wheel speed and hence there is a possibility of torsional oscillations influencing the cutting force in an oscillatory fashion. These oscillating tangential cutting also drive the transverse mode of the machine frame and thus will modify the chatter characteristics of the machine.

Model

To test the hypothesis a mathematical model has been constructed which allows the stability boundary to be determined under various conditions. The grinding force model utilised is that proposed by Chui and Malkin (1993). When sliding and ploughing are ignored the tangential grinding force can be described as:

Figure 1 - System Schematic



where V_w and V_g are the workpiece and grinding wheel surface speeds respectively, b is the width of cut (width of grinding wheel face) and d is the feed per revolution of the workpiece. The constant u_c is termed the 'specific grinding energy'. A schematic of the model used is shown in Figure 1.

$$P_1 = \frac{V_g}{u_c V_w b d}$$

(1)

The solution is progressed by assuming the system to be linear and that all degrees of freedom oscillate with an exponentially modulated sinusoid.

$$x_1(t) = X_1 e^{st}, \text{ where } s = \sigma + j\omega. \quad (2)$$

To locate the neutral stability condition which describes the stability boundary, the real part of the exponential is set to zero, thus imposing neither growth

An experimental result of a 2mm positioning is shown in Figure 6 and Figure 7, it shows a rapid (settling time < 0.5 sec, almost about 0.4 sec), small overshoot (< 0.2 um) and precise (steady state error < 30 nm) positioning is fulfilled. The analysis in this paper and the experimental result show that by combining mechanism design, lubrication and controller design techniques to achieve more robust, rapid, smaller overshoot and precise positioning are expected.

V. Experimental Results, Discussions and Conclusions

we can move the stage into the default settling range(destination position ± 30 nm) rapidly. And its control structure is the same as in Figure 5.

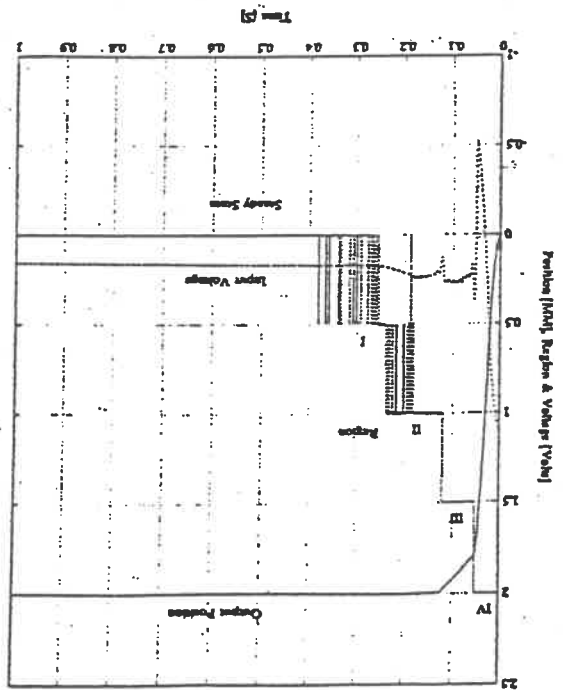


Figure 6. 2 mm positioning under proposed variable structure control algorithm.

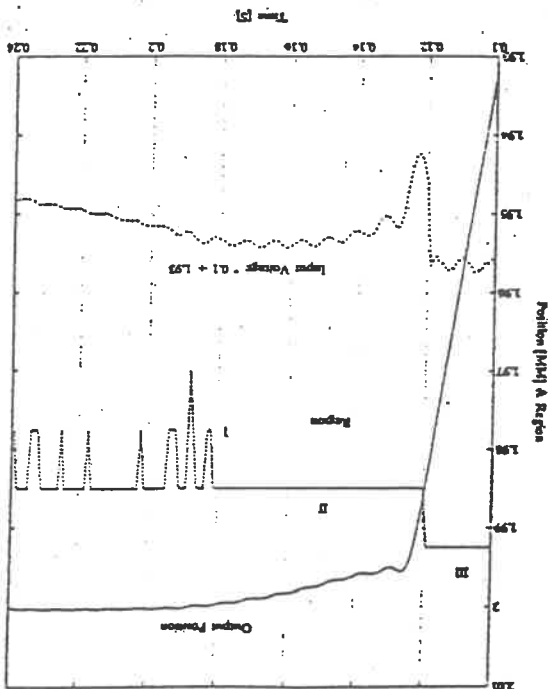


Figure 7. Final positioning condition of 2 mm positioning under proposed variable structure control algorithm.

1. Brian Armstrong-Helouvy, Pierre Dupont and Carlos Canudas De Wit, "A Survey of Models, Analysis Tools and Compensation Methods for the Control of Machines with Friction," *Automatica*, 1994, Vol.30, No.7, pp.1083-1138.
2. Shigeru Futami, Akihito Furutani and Shochiro Yoshida, "Nanometer Positioning and Its Micro-dynamics," *Nanotechnology*, 1990, 1 (1), pp.31-37.
3. P. I. Ro and P. I. Hubbel, "Model Reference Adaptive Control of Dual-Mode Micro/Macro Dynamics of Ball Screws for Nanometer Motion," *ASME Journal of Dynamic Systems, Measurement, and Control*, 1993, Vol.115, pp.103-108.
4. Ogata, *Modern Control Engineering*, Prentice-Hall International Editions, 1990.

VI. References

not decay. Space does not allow a detailed solution to be presented here. Details can be found in Entwistle and Stone (1996). Some of the significant predictions of the model will be outlined here.

The major parameters of interest are the natural torsional frequencies of the workpiece and the grinding wheel relative to that of the transverse mode. These were varied over a large range and found to have a profound affect when in the vicinity of the transverse mode frequency. In what follows, the ratios of the natural frequencies of the torsional degrees of freedom to that of the transverse mode are labelled as F_w and F_g . Only plunge grinding will be considered here.

It is conventional to plot the stability boundary, characterised as the width of cut which gives neutral stability, against the workpiece rotational speed. While usually plotted on linear coordinates for turning and milling, the workpiece speed range for cylindrical grinding is very wide and therefore logarithmic axes are utilised in this work.

Simulations

In order to obtain a benchmark for comparison, the initial simulation is for the situation where both torsional degrees of freedom are inactive. Figure 2 shows the stability boundary predicted in this case.

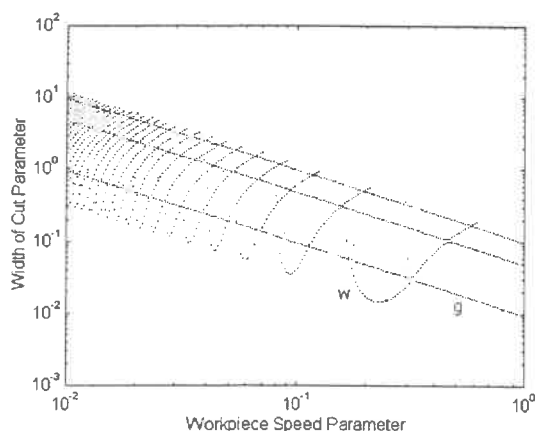


Figure 2 - Stability Plot Without Torsional Effects

Each dot on the plot represents a neutral stability condition, limited to frequencies which are up to twice the transverse natural frequency. The 'straight line' boundaries (labelled 'g') represent the situation where lobes wear on the grinding wheel while the 'U' shaped lobes (labelled 'w') is the situation where lobes wear on

the workpiece. This plot is well known from previously published work.

The direction of the motion of the transverse mode can be modelled at any angle and clearly such motion will affect the surface speeds. Figure 3 shows the stability boundary for the same parameters as used for Figure 2 excepting that $\phi \neq 0$. At lower workpiece speeds the stability boundary is improved and furthermore, below some workpiece speed, the system is unconditionally stable.

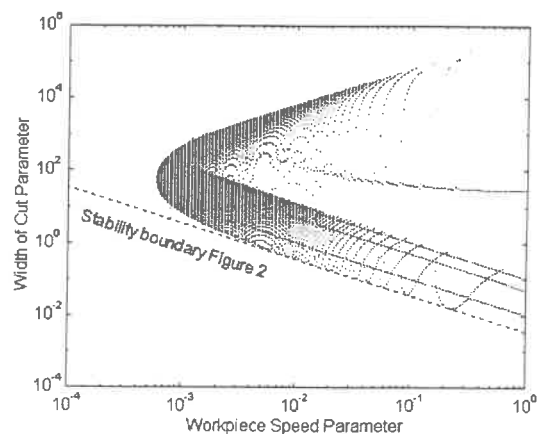


Figure 3 - Introduction of Transverse Mode Angle ϕ

The model is now modified to introduce torsional flexibility on the workpiece mounting and drive system. For illustration purposes the natural torsional frequency of the workpiece was chosen to be 90% of the transverse mode natural frequency ($F_w = 0.9$). Figure 4 is the stability chart for this situation.

It is seen that the introduction of a workpiece torsional flexibility has a similar affect as that due to the transverse mode angle excepting it is a much more dramatic modification. At low workpiece speeds, the model predicts unconditional stability while at higher speeds the boundary is improved. It is noted that the effect is a function of the feed rate; the higher the rate, the further the boundary moves to the right.

The next simulation removes the workpiece flexibility and imparts flexibility to the grinding wheel drive. For illustration purposes the natural torsional frequency of the grinding wheel was chosen to be 110% of the transverse mode natural frequency ($F_g = 1.1$). Figure 5 shows the result.

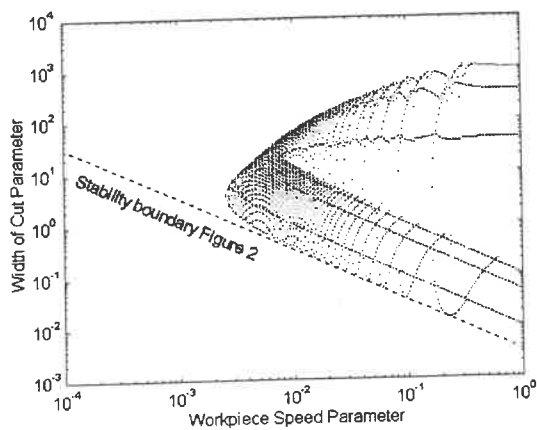


Figure 4 - Introduction of Workpiece Torsional Mode

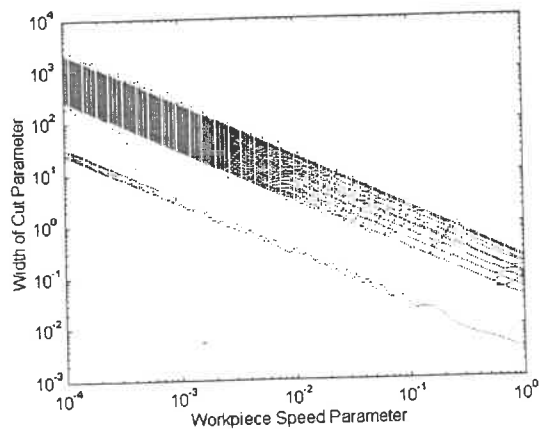


Figure 5 - Introduction of Wheel Torsional Mode

Clearly, if the natural frequency of the grinding wheel and its drive system is near to that of the transverse mode, the stability of the machine can be adversely affected.

The simultaneous application of both workpiece and grinding wheel torsional degrees of freedom results in the stability boundary shown in Figure 6.

The low speed unconditional stability 'cut-off' is reintroduced but above that speed the grinding wheel torsional mode dominates and lowers the stability boundary.

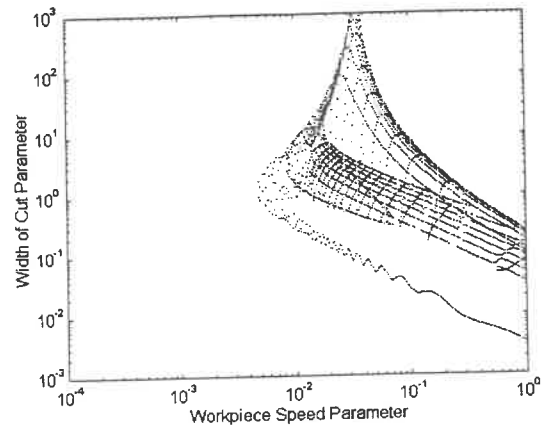


Figure 6 - Introduction of Workpiece and Wheel Torsional Modes

Growth Rates

Many grinding operations take place under unstable conditions but the growth rates can be sufficiently low to permit a considerable amount of useful grinding to be performed prior to the amplitude reaching unacceptable limits. Hence it is always useful to review the predicted growth rates in conjunction with the stability boundaries.

Figure 7 shows the growth rate parameter as a function of the width of cut parameter for the system parameters used to generate Figure 2. The rotational speed parameter chosen here is 0.03.

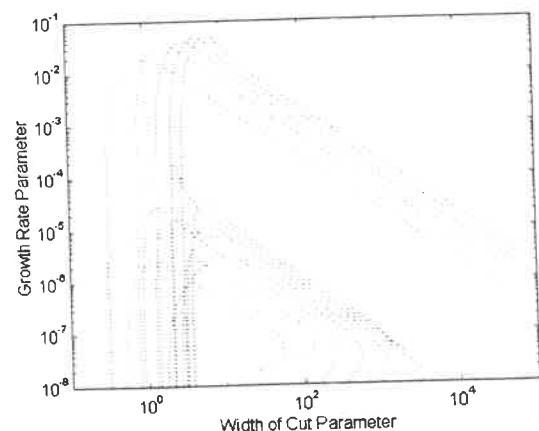


Figure 7 - Growth Rates Without Torsional Effects

The upper curves, which asymptote to zero at the higher workpiece speeds, represent the workpiece regenerative modes whereas the lower curves, at a

much lower growth rates, are associated with the grinding wheel modes which take long periods of time to develop.

In contrast, Figure 8 shows the equivalent plot for the conditions used to generate Figure 4.

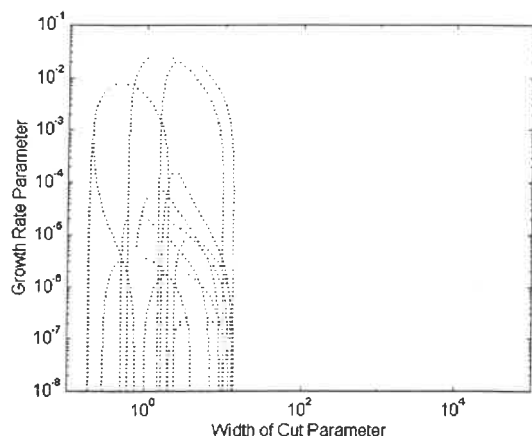


Figure 8 - Growth Rates With Torsional Effects

This last plot indicates that as the width of cut is increased the curves pass through zero growth again rather than asymptoting to zero. Another interpretation of this is that the curves in Figure 4 are closed and so the stability boundary forms a closed curve. Thus, according to the model, regions of stability exist both at low and high widths of cut, with the unstable region existing in between.

Conclusions

A mathematical model of a simplified grinding system has been presented and used to show that workpiece and grinding wheel torsional modes may affect the stability of cylindrical grinding machines. The work-wheel interactions in grinding are very complex and a very simple contact force model has been used. Also

the transverse and rotational systems have been limited to single degrees of freedom. Cognisant of the simplifications made a study with the model has enabled the following broad conclusions to be drawn:

- A transverse mode with a component of its motion in the cutting velocity direction will result in increased stability and unconditional stability below some cutoff workpiece speed.
- The result of a workpiece torsional mode whose natural frequency is in the vicinity of the transverse mode natural frequency is increased stability and unconditional stability below some cutoff workpiece speed.
- When a grinding wheel torsional mode natural frequency is in the vicinity of the transverse mode natural frequency a grinding wheel torsional mode governs stability and is detrimental by lowering the stability boundary.
- If a grinding wheel torsional mode natural frequency is in the vicinity of the transverse mode natural frequency, its detrimental effect can be ameliorated by the introduction of a workpiece torsional mode whose natural frequency is in the vicinity of the transverse mode.
- The growth rates of developing chatter for widths of cut in excess of the stability boundary are modified and, in some cases, reduced to zero by the introduction, and tuning of, torsional degrees of freedom.

References

- Chiu, N., and Malkin, S., 1993, "Computer Simulations for Cylindrical Plunge Grinding" *Annals of the CIRP*. Vol. 42 pp.383-387.
- Entwistle, R.D., and Stone, B.J., 1996, "The Effect of Torsional Vibration on Chatter in Grinding". Submitted to *ASME Journal of Manufacturing Engineering and Science*.

CLOSED-LOOP CONTROL OF A MAGNETIC SERVO-LEVITATED FAST-TOOL SERVO SYSTEM

Hector M. Gutierrez

Department of Electrical and Computer Engineering

Paul I. Ro

Department of Mechanical and Aerospace Engineering

Precision Engineering Center
North Carolina State University
Raleigh, NC

INTRODUCTION

Magnetic Servo Levitation (MSL) [1,7] has been proposed as a method to drive fast-tool servo systems. It represents a promising but very challenging alternative to compensate for errors that can not be adequately handled by machine-tool positioning systems, and to increase the versatility of a given machining process (e.g. to achieve non-rotationally symmetric features in diamond turning).

While MSL could provide an adequate bandwidth within a range of motion twenty times larger than a piezoelectric-based fast-tool servo, there are a number of fundamental issues that make implementation difficult. A long-range MSL-based system is open loop unstable and highly nonlinear, and its dynamics can not be properly represented by simplified models. It is important to emphasize the need of both a highly accurate model and a robust control law if nanometer resolution is expected.

A novel parametric model of the electro-magneto-mechanical coupled system has been developed to describe the behavior of the system over a wide range of motion and frequency, as opposed to more conventional attractive force small perturbation models. It also provides a convenient way of devising a feedback linearizing control scheme that eliminates the need for high biasing currents present in most magnetic bearing systems. This paper presents a feedback-linearized controller coupled with a Kalman filter as a first approach to solve the problem. Good tracking performance has been found both in simulation and experiments.

ELECTROMAGNET DESIGN

The actuator had to be designed avoiding unnecessary mass that could reduce the natural frequency of the system, without sacrificing structural stiffness. Also, electromagnets can only exert attrac-

tive forces and therefore must be used in pairs. Fig. 1 depicts the design concept chosen by the authors.

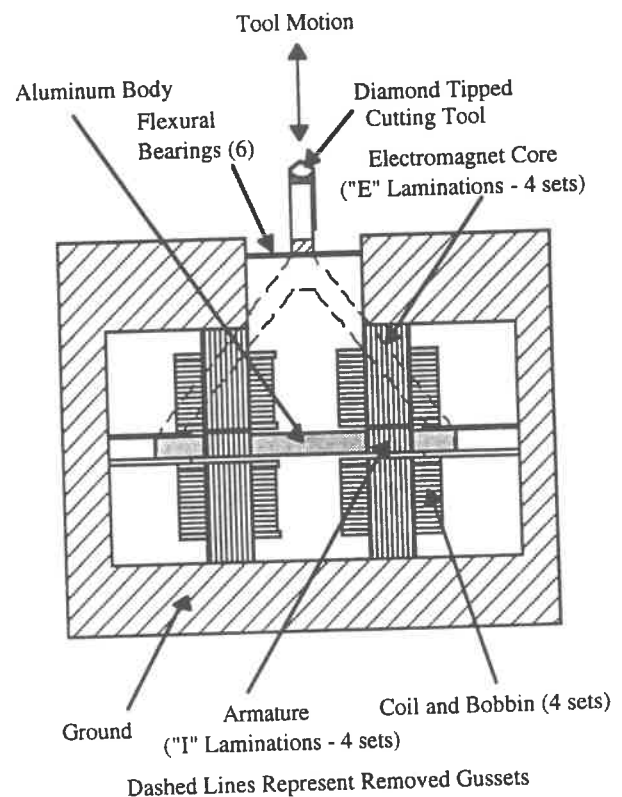


FIGURE 1: ACTUATOR SCHEMATIC

Flexures should be stiff enough to keep the natural frequency of the system above the desired bandwidth, but compliant enough to allow the electromagnets to pull against them and achieve the desired range of motion. This design allows for a tradeoff of bandwidth for range if necessary.

MATHEMATICAL MODELS OF MAGNETIC LEVITATION.

A long range of motion can not be properly described as small perturbations of a nominal trajectory. An adequate model must be accurate over a wide range of displacements and frequencies.

Eletromechanical systems driven by magnetic forces, such as magnetic bearings or micro-positioning stages [9], typically have very small range of motion (in the order of a few microns) and small bandwidth (in the order of a few Hertz). For this reason, simple mathematical models of the magnetic force seem to work reasonably well, for instance:

$$F_{mag} = K \frac{i^2}{g^2}$$

where i is coil current, g is air gap and K some constant to be determined. Furthermore, some of these systems operate at high bias current levels, thus allowing the use of Jacobian linearization. In the case of a magnetically levitated tool actuator, none of these assumptions hold true. This prompts for the development of a different modeling strategy.

A mathematical model can be formulated in such a way that input non-linearities are separable and additive:

$$a_1 \frac{d^2 x}{dt^2} + a_2 \frac{dx}{dt} + a_3 x = F_{mag2} - F_{mag1} \quad (1)$$

where x is displacement, a_i are unknown parameters, and each magnetic force F_{mag} can be described as some polynomial of coil current (i) and the corresponding air gap. For instance:

$$F_{mag2} = \sum_{j=1}^n (K_{1j} \frac{i_1^2}{(x_0 - x)^{0.5j}} + K_{2j} \frac{i_2^2}{(x_0 - x)^{0.5j}}) \quad (2)$$

In these equations, x (displacement) is measured from the equilibrium position, x_0 is the nominal air gap, and K_{ij} are parameters to be determined. The problem is to find the unknown coefficients in Eqs. (1) and (2) such that experimental input-output sequences can be fitted in a minimum mean-square error sense. To use linear parametric techniques to find optimal fits for these parameters, each rational term in Eq. (2) can be considered a separate input to the system, in such a way that (1) can be written as:

$$\frac{d}{dt} \begin{bmatrix} x \\ \frac{dx}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{a_3}{a_1} & -\frac{a_2}{a_1} \end{bmatrix} \begin{bmatrix} x \\ \frac{dx}{dt} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{11} & K_{21} & \dots & K_{4n} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_{4n} \end{bmatrix} \quad (3)$$

The techniques used to estimate the parameters involved are called prediction error methods [8], and are based on the minimization of a cost function which is the sum of the squared prediction errors. Previous knowledge or "reasonable guesses" of the parameters of the dynamic system (e.g. effective mass, estimated spring stiffness) significantly improves speed of convergence and minimizes the chance of hitting local minima since it provides initial estimates to the algorithm. These methods converge to a unique set of parameters and therefore model the system as time-invariant. More details and experimental results can be found in [10].

CONTROLLER DESIGN FOR MAGNETIC LEVITATION BASED ON FEEDBACK LINEARIZATION.

The model structure (Eqs. 2 and 3) is particularly well suited for feedback linearization as described on [2]. This is an approach to nonlinear controller design based on algebraically transforming a non-linear system into a linear one, so that linear control techniques can be applied. This differs entirely from conventional linearization (i.e., Jacobian linearization) in that feedback linearization is achieved by exact state transformations and feedback, rather than by linear approximations of the dynamics. For this reason, this technique is also known as "exact linearization" [3]. Feedback linearization does not require the system to operate around some nominal trajectory nor must the input signal be "small" compared to the input bias level. The MSL fast-tool servo can be feedback-linearized by the following input coordinate transformation:

$$\begin{aligned} \frac{dX}{dt} &= AX + f(X, i_1, i_2) \Leftrightarrow \\ \frac{dX}{dt} &= AX + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \end{aligned} \quad (4)$$

where $f(X, U)$ represents the input non-linearity described by Eq. (2). The problem becomes a standard linear controller design problem, and the performance is limited by the range of the input space $\langle v_1, v_2 \rangle$ over which the inverse map $i = f^{-1}(X, v)$ is well defined. To further clarify this point, consider again Eq. (1) and (2). Terms containing input currents can be factored as follows:

$$\begin{aligned} a_1 \frac{d^2 x}{dt^2} + a_2 \frac{dx}{dt} + a_3 x &= \\ K_1(x)i_1 + K_2(x)i_1^2 + K_3(x)i_2 + K_4(x)i_2^2 \end{aligned} \quad (5)$$

The following input coordinate transformation becomes obviously convenient:

$$\begin{aligned} v_1 &= K_3(x)i_1 + K_4(x)i_1^2 \\ v_2 &= K_1(x)i_2 + K_2(x)i_2^2 \end{aligned} \quad (6)$$

which renders the system linear as in Eq. (4). This transformation is valid subject to the existence of the corresponding inverse maps, i.e.:

$$\begin{aligned} i_1 &= \sqrt{\frac{1}{K_4(x)} \left(v_1 + \frac{K_3^2(x)}{4K_4(x)} \right) - \frac{K_3(x)}{2K_4(x)}} \\ i_2 &= \sqrt{\frac{1}{K_2(x)} \left(v_2 + \frac{K_1^2(x)}{4K_2(x)} \right) - \frac{K_1(x)}{2K_2(x)}} \end{aligned} \quad (7)$$

Once the system is linearized, controller design is carried out based on a linear quadratic regulator (LQR), using a reduced-order Kalman filter to estimate the non-measurable states [4].

A singular value plot on the frequency domain shows that the system has inherently poor tracking capabilities. For this reason, the integral of the tracking error was added as an additional state to take advantage of the regulatory capabilities of the LQR design [5,6].

The feedback sensor used was an infrared optical probe. Tests were carried out on a smaller fixture, with a natural frequency of 190 Hz; in all plots, command signal is plotted in dashed line, measured response is shown in solid line.

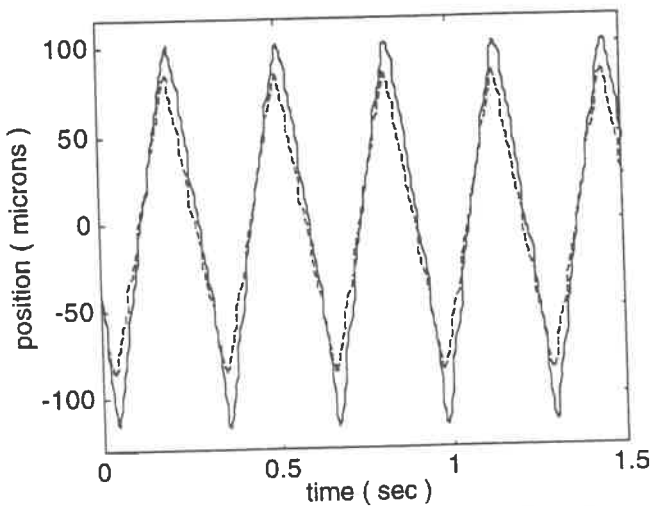


FIGURE 2. RESPONSE TO RAMP COMMAND

Figures 2 and 3 illustrate the tracking capabilities of the augmented LQR compensator. Tracking performance, however, degrades when either range or frequency bounds are exceeded. Since feedback linearization is sensitive to sampling rate, this deviation was primarily attributed to limitations in

hardware used for these tests, in which closed-loop rate was limited to 2.5 kHz.

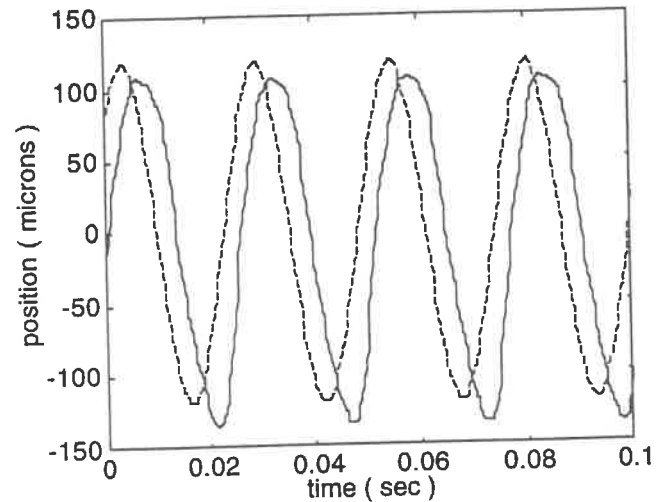


FIGURE 3. RESPONSE TO SINE WAVE COMMAND (40 Hz)

To further illustrate this, the closed loop system was simulated at a closed loop rate of 50 kHz, assuming a computational delay of 10 μ sec (half the sampling period). Responses to a 400 micron step command are shown in Figure 4. This illustrates one key feature of exact linearization: the difference in response between actual plant and linearized system asymptotically decreases as the sampling time grows.

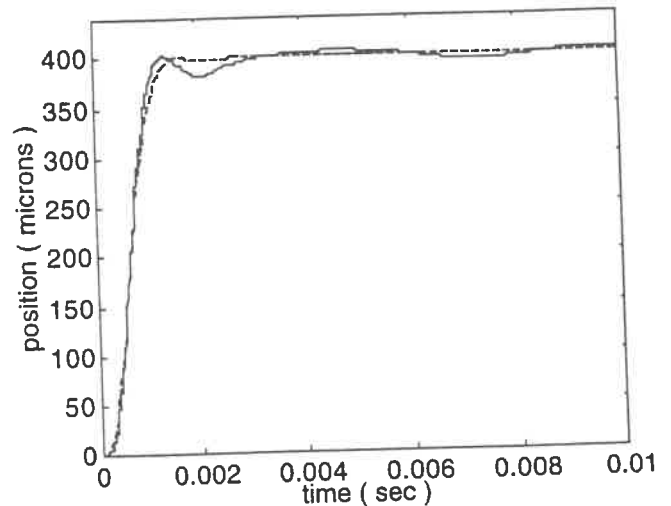


FIGURE 4. STEP RESPONSE AT 50 KHz (linearized model: dashed, nonlinear plant: solid)

At the current sampling rate both actual plant and nonlinear model fail to track the reference command beyond some range and frequency. To examine this assumption, input coil currents were recorded while operating the system in closed loop. Current coil data

was later sent through a simulation network corresponding to Eq. (1) and (2) in order to evaluate the model ability to reproduce the behavior of the nonlinear plant. Results showed good agreement between simulated and measured response.

Other plausible factors that degrade overall performance are inherent model limitations, amplifier bandwidth and sensor noise. Domain boundaries imposed by the inverse maps (Eq. 7) and amplifier bandwidth (frequency band over which the power amplifiers behave as ideal current sources) may also be affecting overall closed loop performance.

Experimental frequency response of the closed loop system is shown in Figure 5. Gain at frequencies below the resonance shows the range at which good tracking can be achieved with the current setup. This should significantly improve (get closer to 0 dB for frequencies below the resonance) at higher closed loop rates.

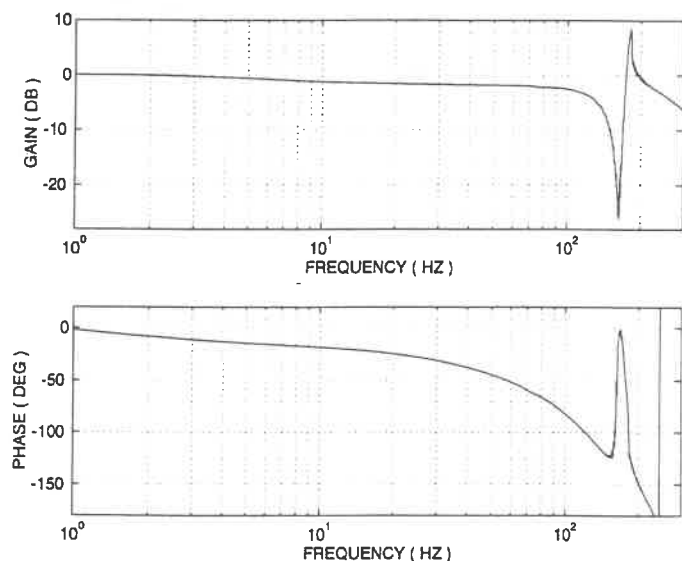


FIGURE 5. CLOSED-LOOP FREQUENCY RESPONSE

CONCLUSIONS

Magnetic Servo Levitation represents a promising alternative to long range tool actuation while keeping an adequate bandwidth.

A novel mathematical model of magnetic levitation has been presented and used for controller design. The model has been experimentally tested both by simulating actual data and by closed loop tests. It is important to notice that an adequate model is essential to control an electromagnetic actuator since there is a highly nonlinear map between error signal and control action (coil currents).

The augmented LQR compensator has shown good tracking capabilities, although within a limited range and bandwidth. A substantial increase in closed loop rate may yield significant performance improvement.

Several other effects involved in the use of the magnetically levitated actuator are yet to be investigated. Inclusion of adaptive or learning schemes as a way of handling time varying effects and/or unmodelled dynamics represents an interesting path for future research.

ACKNOWLEDGEMENTS

Authors gratefully acknowledge David Trumper from MIT for his help on electromagnet design at an early stage of this project. This work was partially supported by ONR (Contract Number N00014-94-1-0235).

REFERENCES

- [1] Tsuda, M. and T. Higuchi, "Design and Control of Magnetic Servo Levitation", *Report of the Institute of Industrial Science*, The University of Tokyo, March 1992, Vol. 37, No. 2, pp. 137-205.
- [2] Slotine, J. J. and Li, W. : *Applied Nonlinear Control*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [3] Khalil, H. : *Nonlinear Systems*. Macmillan Publishing Company, New York, NY 1992.
- [4]. Phillips, C. L. and Nagle, H. T. : *Digital Control System Analysis and Design*. Prentice-Hall, Englewood Cliffs, NJ, 1991.
- [5] Franklin, G. F. and Powell, J. D. : *Digital Control of Dynamic Systems*. Addison-Wesley Publishing Company, Reading, MA, 1980.
- [6] Dorato, P., Abdallah, C. and Cerone, V. : *Linear-Quadratic Control: An Introduction*. Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [7] Higuchi, T., Tsuda, M. and Nakamura, Y.: "Design of Magnetic Servo Levitation for Robot Mechanisms". *Proc. of 20th Int. Symp on Industrial Robots*, 1989.
- [8] Ljung, L.: *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
- [9] Trumper, D. L., Ludwick S.J. and Holmes, M.L.: "Design and Control of a 6-DOF Magnetic/Fluidic Motion Control Stage" *Proceedings of the ASME Int'l Congress and Exposition*, San Francisco, CA, 1995.
- [10] Gutierrez, H. M., Stancil, B. A. and Ro, P. I. : "Design and Modeling of a Magnetically Levitated Fast-Tool Servo System for Precision Turning". *Proceedings of the ASME Int'l Congress and Exposition*, San Francisco, CA, 1995.

Performance evaluator for open architecture NC boards

Minoru Hatakeyama, Shigeru Ueno

Technical Research Institute of
Japan Society for the Promotion of Machine Industry
1-1-12, Hachiman-cho Higashikurume-shi Tokyo, Japan

1. Introduction

The need for an open architecture for NC systems has increased⁽¹⁾⁽²⁾⁽³⁾, and many makers have begun to produce this type. The open architecture NC system has high flexibility and the ability to meet many sophisticated requirements such as adaptive control and intelligent prediction. Due to this advance, machine tool manufacturers today can choose and build the best matching NC system for their own products from various NC suppliers. However it is still difficult to select the best combination for the target machine from the many NC systems, servo amplifiers and other needed equipment because a method of evaluation for these has not yet been established. This research project is to develop a device that can evaluate many types of open architecture NC boards easily and consistently.

One of the major problems facing manufacturers is the need to assemble an NC system and

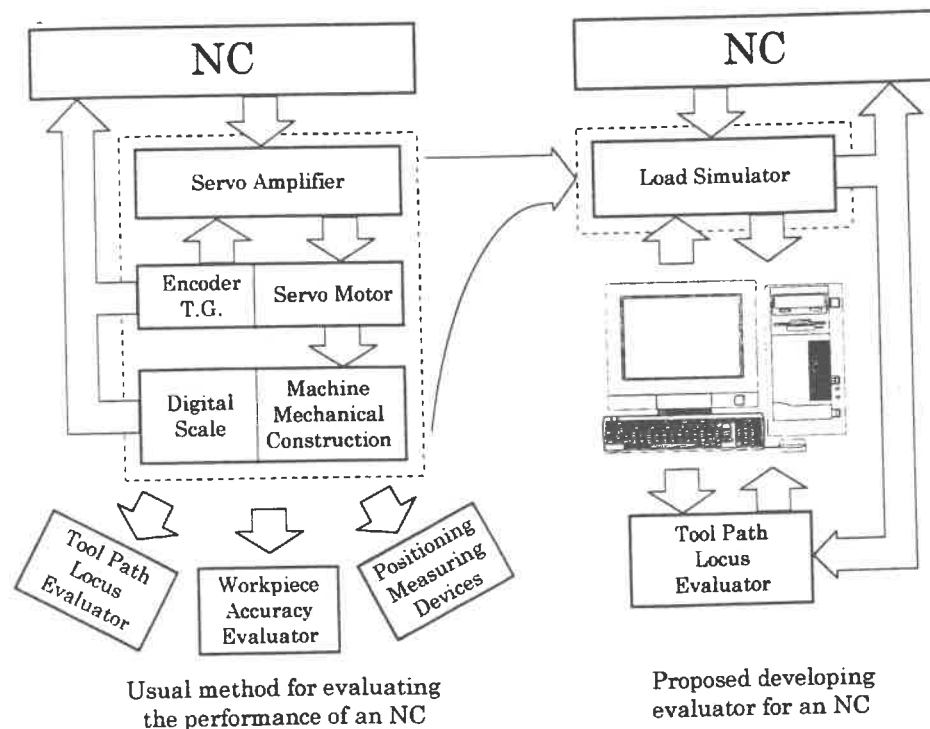


Figure 1 The examples of the method for evaluating a performance of NC

the servo amplifier on the target machine, and to activate it for evaluating the performance of the NC system.

However even this method's results are affected by the influence of the servo amplifier or the servomotor and therefore it is difficult to accurately evaluate the performance of the NC system itself. Also a standard system of evaluation has not yet been established.

2. Design Target and System Construction

An example comparing the present method for evaluating the performance of the NC system and the proposed performance evaluator for an NC system, are shown in Figure 1.

The proposed performance evaluator consists of a digital circuit that can simulate servo amplifiers and servomotors, and a load simulator that can simulate the target machine's mechanical characteristics. Due to the fully digital construction and software algorithm, the simulator can constantly evaluate performance according to each situation. With the digital circuit, the evaluator will be able to be set up various conditions from a personal computer. Information from the NC system is matched in the evaluator and converted to a digital signal. The converted signal is then sent to the various other simulator circuits.

As shown in Figure 2, the simulator consists primarily of four circuits. The first circuit is an IIR (Infinite Impulse Response) filter for linear responses. The second and third circuits are the simulators for nonlinear responses, i.e., the influences of friction that include stick slip motion, and mechanical backlash caused by changing one quadrant to another quadrant on Cartesian

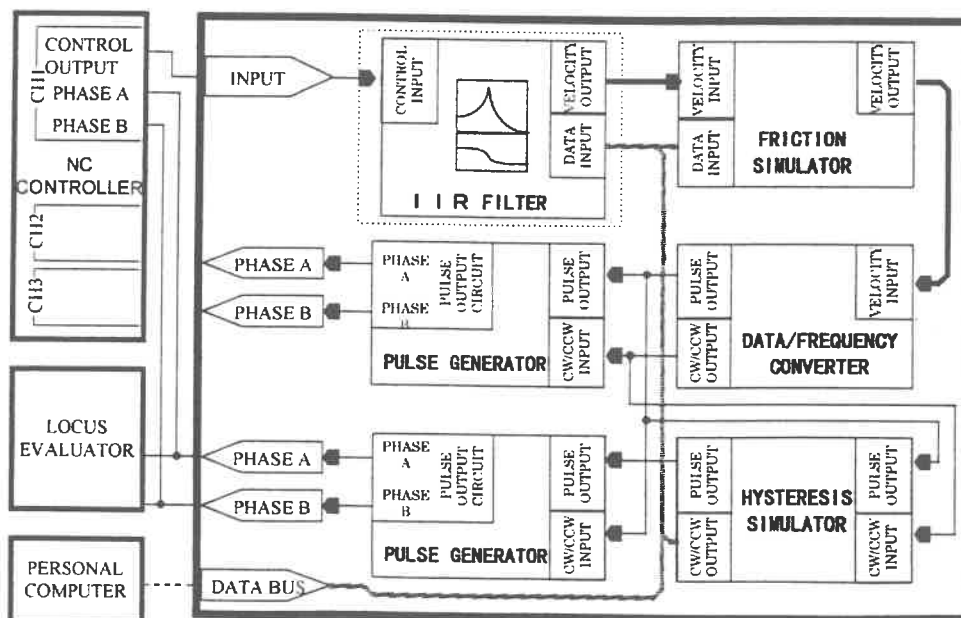


Figure 2 Outline of the Load Simulator

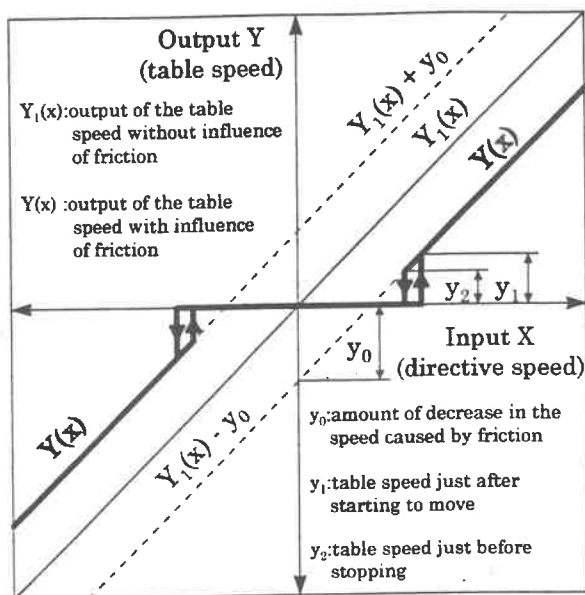


Figure 3 Output speed pattern considering the influence of friction

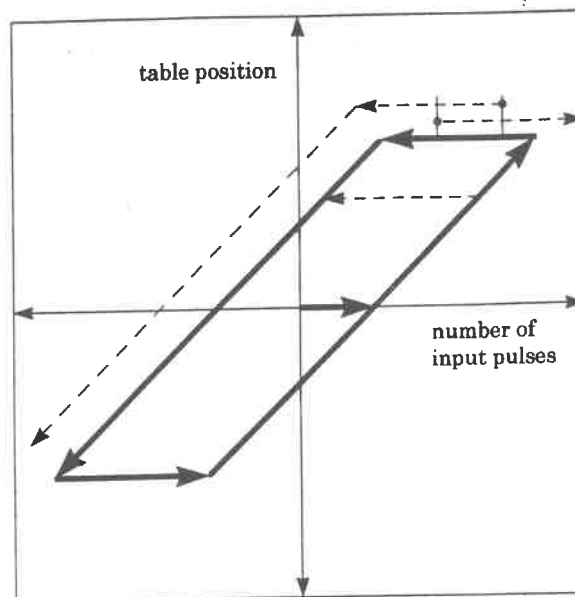


Figure 4 Behavior of the table considering the influence of mechanical backlash

coordinates. The last circuit is a pulse translator for feedback to the NC system. The system can accept AT bus NC boards that can export velocity analog outputs, and evaluation calculations are performed on two axes motions.

3. Kinematic models and its simulated device

One major focus of this study is the simulation of various kinematic behaviors. The mechanical and nonlinear behavior models, are shown in Figure 3 and Figure 4.

The behavior caused by the influence of friction, can become a simple model, so a digital circuit can simulate it in the friction simulator. In this circuit, the amount of decrease in the table speed by friction, the speed just after starting to move and the speed just before stopping can be set up independently. The combination of this circuit's effect and the delay response with the IIR filter in front of it, will be able to simulate the movement of stick and slip.

Figure 5 shows the outline of the circuit diagram. This circuit consists of comparators, data selectors and an adder. The adder computes the speed that is decreased by the friction. The comparators and the data selectors select the threshold value that is matched with the order of velocity. The hysteresis simulator circuit that simulates the influence of mechanical backlash, does not only cut the suitable pulses for it but traces the table position with a counter so it can prevent the table position from shifting in the backlash area.

After these procedures are performed, velocity output is calculated, and through the data/frequency converter and pulse generator, quadrature position pulses are generated as the

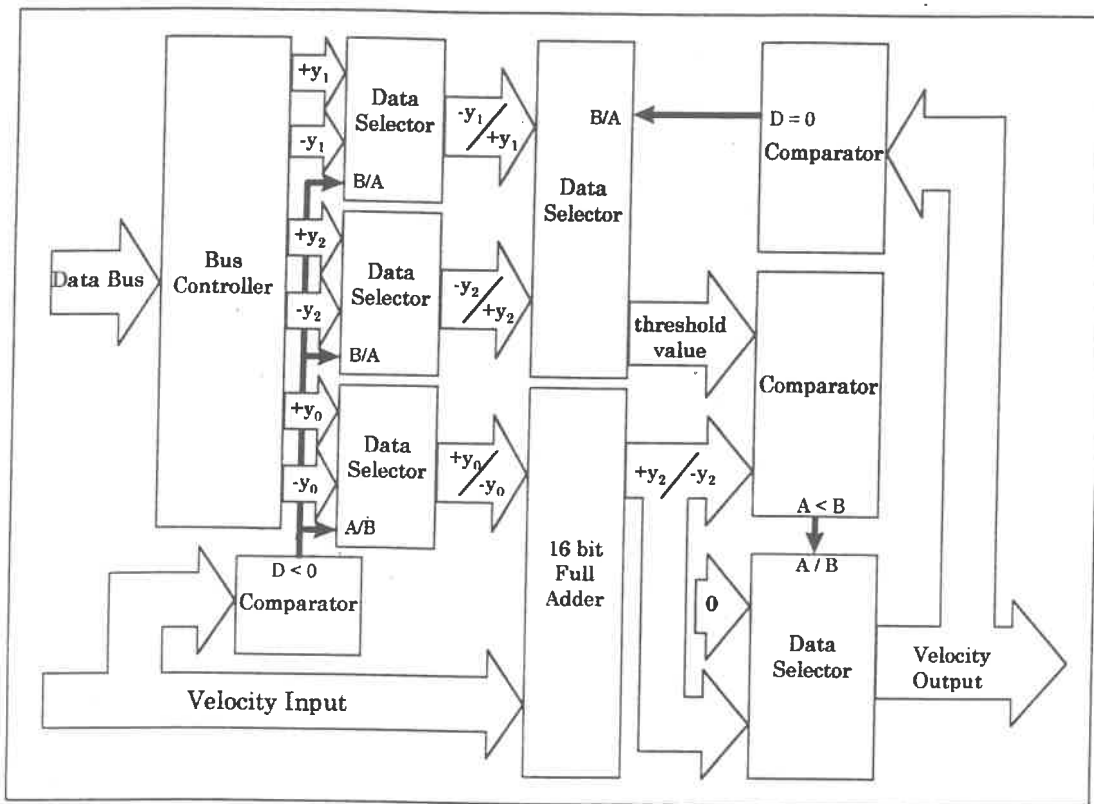


Figure 5 Friction Circuit

machine position. This information quality is evaluated by the digital locus evaluator that can calculate the difference between control position demands and simulated machine positions.

4. Future focus

The proposed system can evaluate various NC boards without actual machine hardware simulating the same specified conditions. The obtained output can be used for the comparison of different NC boards' performance, and will also be useful for the improvement of NC boards.

5. Reference

- (1) Requirements of Open, Modular Architecture Controllers for Applications in the Automotive Industry, Version 1.1, (1994), Chrysler, Ford, GM
- (2) Open Architecture Controls for Precision Machine Tools, Bradley N. Damazo, Charles W. Yang, Robert J. Gavin, National Institute of Standards and Technology, the Proceedings of American Society for Precision Engineering Spring Topical Meeting on Mechanisms and Controls for Ultraprecision Motion, April 6-8, 1994, Tucson, AZ, pp. 26-31.
- (3) OSEC Architecture Version 2.0 Pre-release, August 1996, OSE study members

Fabrication of Surface Perturbations on Inertial Confinement Fusion Targets

B. H. Jared, T. A. Dow, K. P. Garrard and G. M. Moorefield III
Precision Engineering Center, North Carolina State University, Raleigh, NC

C. Barnes, R. D. Day, D. J. Hatch, L. J. Salzer and G. Rivera
Los Alamos National Laboratory, Los Alamos, NM

Key Words: Inertial Confinement Fusion, diamond turning, fast tool servomechanism

Introduction

The production of inertial confinement fusion (ICF) targets for the National Ignition Facility is an important goal being pursued at Los Alamos National Laboratory (LANL). The geometry of ICF targets must be very well controlled as surface perturbations can create hydrodynamic instabilities when the targets are imploded during the fusion process. Targets with known surface features are currently being fabricated and tested in an attempt to better understand the impact of surface perturbations in ICF. Experimental data collected using targets with known surface features is also being compared with theoretical data generated by computer models.

Surface perturbations on flat target surfaces were initially produced by LANL using a standard diamond turning machine. Surface features have also been machined onto cylindrical ICF targets using both a precision fly-cutter mounted onto a diamond turning machine (DTM), and a piezoelectrically actuated tool which follows a reference surface on a rotating cam. Since each of these methods have limitations, an alternate target fabrication process has been developed at the Precision Engineering Center (PEC) which utilizes a fast tool servomechanism (FTS) mounted onto a diamond turning machine. The FTS system will give LANL the ability to produce more complex surface geometries with greater precision and a substantially reduced machining time than was previously possible.

Target Specifications

The manufacture of ICF targets requires a high degree of machining accuracy and precision. The cylindrical targets fabricated by LANL involve machining surface features around the circumference of 430 μm diameter cylindrical targets which are 430 μm in length. The current

fabrication techniques at LANL produce surface features which resemble a $0.5\text{ }\mu\text{m}$ amplitude sine wave pattern with a frequency of 10 waves per target revolution. Additional ICF tests specify the production of similar sized targets with up to 50 sine waves per revolution. LANL also has plans to produce spherical targets by machining $0.5\text{ }\mu\text{m}$ amplitude spherical harmonics on the outside diameter of $430\text{ }\mu\text{m}$ diameter hemispheres. Spherical targets cannot be produced using current LANL fabrication techniques, but should be machinable using a FTS. The desired figure accuracy of the surface perturbations on both the cylindrical and the spherical targets is 10 nm peak to peak.

FTS Design

The FTS built for LANL, as seen in Figures 1 and 2, uses the same basic concepts found in the FTS developed by Falter [1]. It was specifically designed, however, to meet the specifications

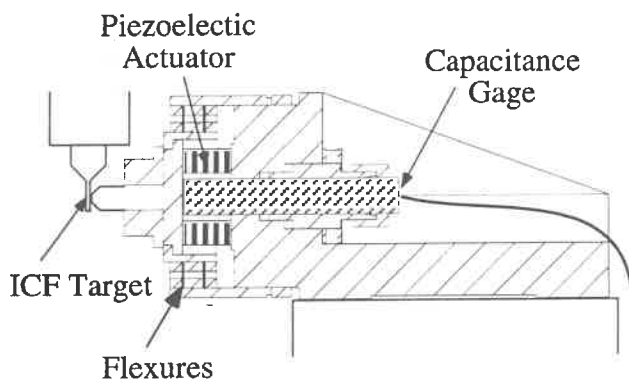


Figure 1: FTS Cross Section

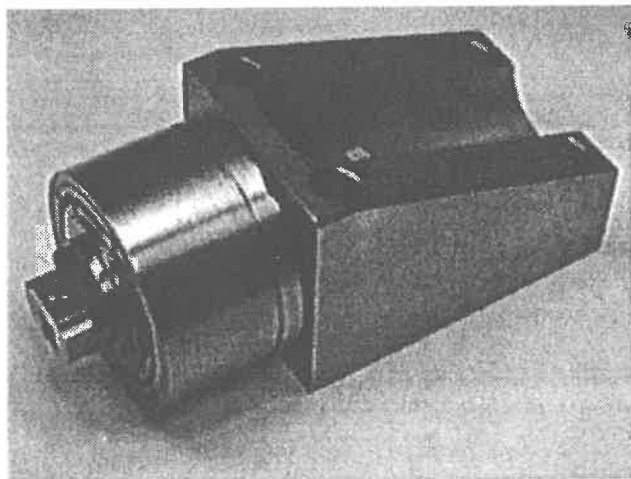


Figure 2: FTS

required for the fabrication of ICF targets and for the DTM hardware utilized by LANL. The FTS uses an EDO E400P-0.5 piezoelectric actuator to provide $11.4\text{ }\mu\text{m}$ of tool motion. This tool motion is measured using a Lion Precision capacitance gage system with a PX405HC probe and a DMT10R probe driver. The Lion Precision system provides a measurement resolution of 1 nm at a bandwidth of 21 kHz . The moving components of the FTS were optimized such that their natural frequency approached 10 kHz . This was done by modeling the FTS in Pro/ENGINEER and then performing finite element analyses on the model using ANSYS. Since ICF targets are currently fabricated at LANL using a Moore M-18 Aspheric Generator, the FTS was also designed to mount to the side of the tool post of the M-18 so that the tip of its diamond tool could be easily positioned at the center of the rotary table in the horizontal plane and level with the spindle axis in the vertical plane.

FTS Control

A FTS cannot generate the tool motions necessary to machine surface perturbations onto ICF targets while operating in an open loop mode due to hysteresis in the motion of the piezoelectric actuator and due to the necessity of coordinating the motion of the FTS with the motion of the spindle, axes and rotary table of the M-18. Since the M-18's Aerotech Unidex 31 (U31) controller does not provide the software and hardware configurations necessary to implement a FTS system, modifications were made to the U31 controller in the form of a separate control program for the FTS. The major hardware modifications to the U31 involved installing a Pentek 4284 DSP board and a Pentek 4242 Analog I/O board into its VME-bus. This allows the DSP board to perform the computations necessary to execute the FTS's PID control algorithm at an update rate of 30 kHz. Modifications were also made to the U31 program code to allow it and the FTS control program to operate concurrently.

The design of the control algorithm for the FTS involved optimizing the PID control gains using

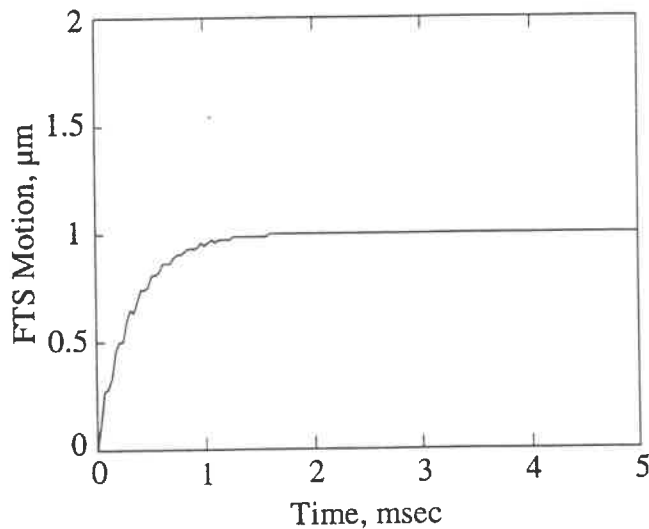


Figure 3: Simulated Step Response

MATLAB simulations. An experimental gain plot of the FTS was obtained by measuring the response of the servo to a 50 V amplitude swept sine wave input with a maximum frequency of 15 kHz. A mathematical transfer function was fit to the experimental data and then utilized in the controller simulations. The PID control gains were varied through trial and error until the simulated response to a 1 μm step input was optimized, as seen in Figure 3. The simulation predicts that the FTS will exhibit a settling time of about 1 msec with zero overshoot and zero steady state error for a 1 μm step input.

FTS Testing

Once the PID control gains were obtained, initial tests were performed on the FTS to determine its response characteristics and to improve upon them. These tests were performed on a Rank Pneumo ASG 2500 located at the PEC. The ASG 2500 uses a controller designed at the PEC which only provides a 10 kHz sampling rate. This controller, however, has the hardware and

software necessary to implement a FTS system. Thus, initial testing of the FTS was performed with a minimal amount of controller modification. The FTS was commanded to follow a 1 μm

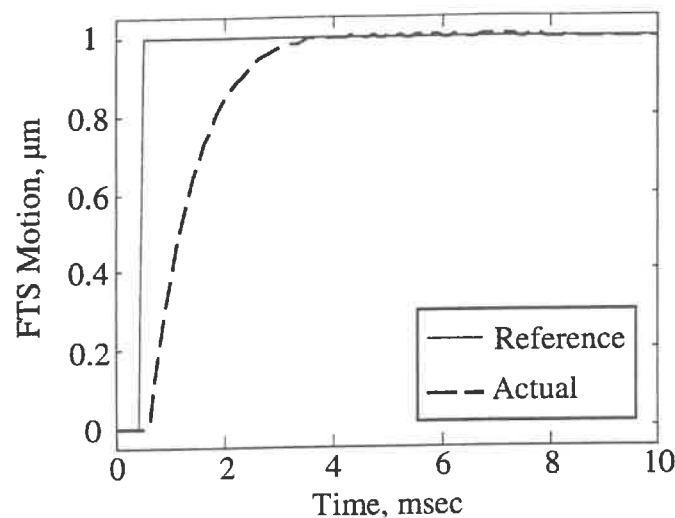


Figure 4: FTS Controlled Step Response

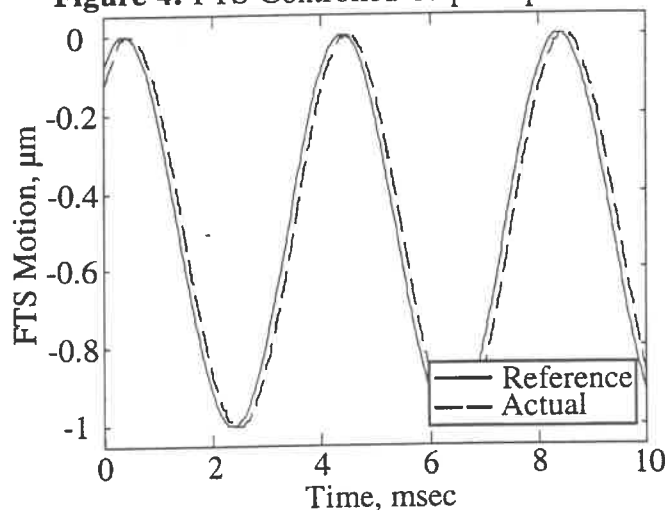


Figure 5: FTS Sine Wave Response

square wave at 2 Hz to determine its response to a step input. Figure 4 shows the response of the FTS after the PID gains were optimized. It demonstrated a settling time of almost 4 msec with an overshoot of less than 0.6% and a steady state error of less than 0.1%. The FTS was also commanded to follow a 0.5 μm amplitude sine wave at 33 Hz. The FTS, as seen in Figure 5, demonstrated ± 80 nm of following error. Much of this error, however, was associated with the phase response of the FTS as the error motion from a true sine wave geometry was only ± 5 nm. Increasing the sampling rate of the controller to 30 kHz will greatly improve the phase response of the FTS and thus significantly reduce the following error.

Conclusions

A FTS has been designed and built for LANL. Initial tests on the system indicate that it will generate surface perturbations which meet current ICF target specifications. The FTS,

however, must still be interfaced with the Aerotech Unidex 31 controller. Once the FTS can be controlled by the U31 controller, it will be transferred to LANL as ICF targets will be machined with the M-18 to determine the response of the FTS system under machining conditions.

References

- [1] Falter, Peter. "Diamond Turning of Nonrotationally Symmetric Surfaces," PhD Thesis. North Carolina State University. 1990.

AUTONOMOUS CUTTING PARAMETER REGULATION BY USING AN ADAPTIVE MODELING AND GENETIC ALGORITHMS

Tae Jo Ko, Hee Sool Kim
Department of Mechanical Engineering, Yeungnam University
Gyeongsan, Kyoungbuk 712-749 KOREA

ABSTRACT

In this research, a turning process was modeled adaptively by a back-propagation neural network with iterative learning method, and cutting parameters of the process model were optimized through a Genetic Algorithms(GAs). Some constraints were given to the input conditions and the process outputs for the desired surface integrity and to protect a machine tool. Such a constrained problem could be solved by introducing penalty values, which were included in the fitness evaluation of GAs. Experimental results show that the neural network could model the cutting process efficiently, and the cutting conditions such as spindle speed and feed rate were regulated for maximizing material removal rate through the GAs. Therefore the proposed approach can be well applied to the real machining process.

INTRODUCTION

In conventional machining systems, the cutting parameters such as cutting speed, feed rate and depth of cut are determined based on the operator's experience or empirical data hand book. Such systems can not be optimal in the sense of material removal rate since the cutting conditions are selected conservatively considering tool breakage or machine tool damage[1].

In this regard, the adaptive control system that can adjust the operating parameters on-line has been studied with interest. Adaptive control systems are classified into two types: adaptive control of constraints(ACC) and adaptive control with optimization(ACO)[5]. Although there have been considerable researches on the ACO systems[1-4], few of these systems are used in practice. The major problems with such systems have been the difficulties in modeling the cutting process. Although there are well known cutting process models, such models can not reflect the progress changes such as tool wear or any environmental effect. Adaptation of the cutting conditions to the new circumstances means better optimization in view of a greater utilization of the system's capabilities. Moreover, although there have been mathematical process model that can adapt the process changes, a calculus based optimization technique has a chance to give a local optimal value[10]. This is due to that the cutting process is rather complex and nonlinear with respect to materials, cutting conditions, cutting tools, etc.

In this paper, an adaptive optimization method of machining process using an adaptive modeling and GAs is introduced[6-8]. To this end, a multi-layered neural network is used to model the cutting process on-line.[9] The network learns the machining process iteratively using the errors between the sensor and the network outputs without supervisor. Therefore this model can adapt the process change such as the variations of tool wear or change of work-piece geometry. Then GAs are introduced to optimize the cutting parameters under the given constraints such as cutting force, motor current and machine tool vibrations. Consequently cutting parameters are continuously adjusted through the feedrate and spindle override function in NC machine tools. The result shows that a multi-layered neural network learns the process by machining two or three pieces of work-piece(total cutting length 50mm). The cutting parameters such as feed rate and spindle speed are adjusted autonomously to maximize the material removal rate without violation the given process constraints.

ADAPTIVE MODELING OF TURNING PROCESS

The optimization system consists of the engine for modeling the turning process adaptively and the one for optimizing the modeled process. As shown in Fig. 1, an error back-propagation multi-layered neural network is used for modeling, and GAs are introduced for optimization. In the modeling, the u_k of cutting parameters such

as cutting speed, feed rate, and depth of cut, which are initially given by the CAM software, is fed to the machine tool and the neural network simultaneously. The predicted outputs of a neural network may be wrong since the network never learned the process previously. The same cutting conditions given to the CNC machine tool also give the process output such as cutting force, machine tool vibration, and spindle motor current. Therefore there exists an error between the predicted value of network and the output of machining process, and the error is minimized by training the network recursively.

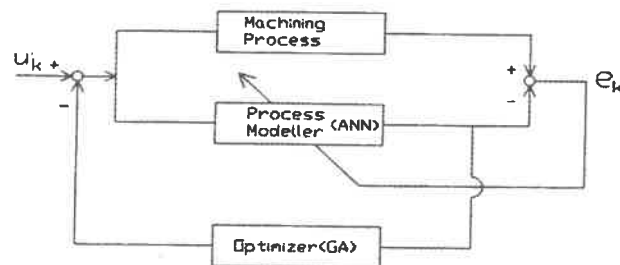


Fig. 1 The overall structure for the process optimizer

Therefore the optimized value determined from the GAs can not be optimum, and new input conditions are generated by feeding back the above value to the u_k of original inputs. The new cutting parameters are given to the network and the machine tool again, and the on-line learning process keeps on running until the error is within tolerance value. The neural network used in the research is multi-layered neural network which has three layered structures, that is input, hidden, and output layer.

OPTIMIZATION OF MACHINING PROCESS

The material removal rate can be represented by multiplying cutting speed, feed rate, and depth of cut as shown Eq. (1).

$$H = f \cdot d \cdot v \quad (1)$$

where f is feed rate, d is depth of cut, and v is cutting velocity. Feed rate and cutting speed are selected as the control parameters since the control technology uses the external input such as feed rate and spindle override function in NC. The change of depth of cut may be difficult in finish machining, therefore the control of depth of cut was omitted.

In Eq. (1), maximum material removal rate can be obtained by choosing maximum allowable values of three input parameters. However, some constraints should be given for assuring surface integrity and protecting a machine tool. That is, the optimization problem of Eq. (1) has some constraints as follows;

Input constraints :

$$f_1 < f < f_2, s_1 < s < s_2 \quad (2)$$

Output constraints

$$F_c < F_{c \max}, I_s < I_{s \max}, A_s < A_{s \max} \quad (3)$$

where $F_{c \max}$, $I_{s \max}$, and $A_{s \max}$ represents the maximum constraint value of main cutting force, spindle motor current, and acceleration signal respectively. f_1, f_2, s_1, s_2 is the minimum and the maximum allowable value for the feed rate and spindle speed respectively.

The above constrained optimization problem can be transformed to unconstrained problem by penalty function technique as shown in Eq. (4). The above constrained optimization problem can be transformed to unconstrained problem by penalty function technique as shown in Equation (4).

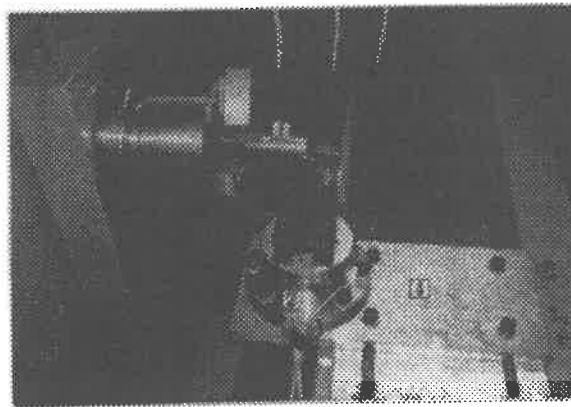
$$\begin{aligned}
 H = & f \cdot d \cdot v \\
 & - R_1 (< f_1 - f >^2 + < s_1 - s >^2 + < f - f_2 >^2 + < s - s_2 >^2) \\
 & - R_2 (< F_c - F_{\max} >^2 + < I_s - I >^2 + < A_s - A >^2)
 \end{aligned} \quad (4)$$

where $<\bullet>$ represents $<g>=g$ when $g \geq 0$, $<g>=0$ when $g < 0$. R determines the relative penalty values imposed on the constrained conditions. These values play a role to reduce the fitness value in the application of GAs. The first term in Equation (4) is fundamental objective function, and the second one is the constraints to the input conditions. The other represents the maximum allowable output values to the cutting force, spindle current, and machine tool vibration. By applying GAs to the Equation (4), the globally optimized cutting conditions can be taken.

EXPERIMENTAL METHOD

Experimental Apparatus

A small CNC turning lathe was used for the experiment. Tool dynamometer, and accelerometer were used for measuring process output. Tool dynamometer(KISTLER 9272) for measuring cutting force signal was mounted on the fixture: fixed on the gang type table as shown in Fig. 2. Instead of accelerometer for measuring vibration signal, piezo electric film(Pennwalt; sensitivity 10mV/g) was attached under the tool holder. The reason is that we can fabricate it as a lower price to commercialize the system. The current signal was directly acquired from the spindle drive unit(A06B-6060-H003#H503), which is the line for the spindle load meter in operating console. Therefore we need not a current sensor or any other type of composition method for three phase current signals. The above signals were digitized through the A/D converter(PCL 812PG) and sampled to the PC. After optimizing the input cutting conditions, the cutting parameters were regulated through the feed rate and the spindle speed override function by a specially made I/O card. The work-piece used in the cutting was carbon steel(SM45C), and the cutting tool is tungsten carbide(TNMG 160412-B20).



(1) Workpiece (2) Tool Dynamometer (3) Piezo Electric Film (4) Bracket
Fig. 2 The Experimental Set Up

Experimental Method

In the experiment, firstly, cutting force signals, vibration signals, and current signals are acquired through the sensors. Then the signals are used for training the neural network. Finally, the GAs optimizes the cutting parameters and new optimized cutting parameters are fed to the feed rate and the spindle speed override function in the CNC(FANUC 0T).

The override function in the CNC system are mainly consisted of feed rate and spindle speed, and which regulates the cutting parameters a little bit to the given initial values. For example, the feed rate uses four bits, and the combination of each bit has different values of override. Therefore, the override in feed rate has sixteen

step from 0% to 150% to the given initial value. On the other hand, the spindle override uses three bits, the range is from 50% to 120%. The functions on the operating console were switched by the control system through the PC.

EXPERIMENTAL RESULTS

Firstly the CNC turning center cut the work-piece without any information about the machining process. Of course, the machine has some constraints which were chosen from operator's experience. The main cutting force should be under 100N, and the acceleration signal be under 2.5g. The maximum current to the spindle motor is 15A. The range of feed rate is from 0mm/rev to 0.75mm/rev, and the spindle speed is from 500rpm to 1200rpm. Two kinds of cutting experiments were carried out; one for the geometric change such as depth of cut, and the other for the progressive tool wear.

Cutting Parameter According to the Change of Depth of Cut

Fig. 3s show the optimized results to the depth of cut of 0.3mm. The initial population in the GAs is 100, the crossover rate is 0.25, and the mutation rate is 0.01. Fig. 3(a) represents the maximum and averaged fitness value during the evolution of 1000 generations. Feed rate variation according to the evolution is Fig. 3(b), and Fig. 3(c) shows the change of spindle speed. As shown in the figures, the cutting parameters converged to the optimal value so fast in case of depth of cut of 0.3mm. The neural network learns the machining process simultaneously as shown in Fig. 3(d), where the error tolerance is 0.01, learning rate is 0.5, and momentum rate is 0.5. The optimized cutting parameter of feed rate and cutting speed is 0.75mm/rev and 1200rpm, respectively, which is the maximum value to the given constraints. Therefore, the maximum cutting conditions do not violate the given constraints to the small depth of cut.

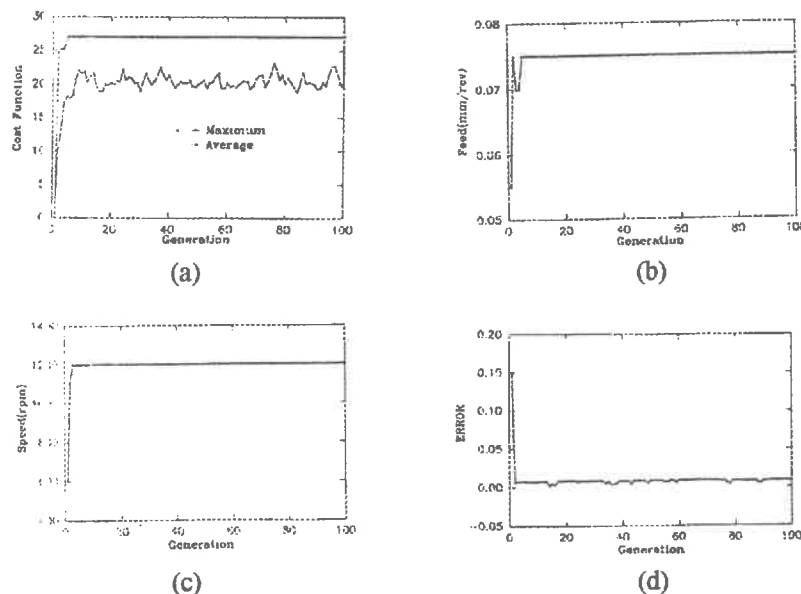


Fig. 3 Results of Genetic Algorithm according to the evolution; (a) Maximum and averaged values of evaluated costs, (b) Variation of feed rate, (c) Variation of cutting speed, (d) Learning errors of a neural network.

Fig. 4s show the optimization results by changing the depth of cut from 0.3mm to 0.6mm. At this time, the neural network trained to the depth of cut of 0.3mm was used again. As shown in Fig. 4(a), there are big oscillation to find the optimum value initially since the network has wrong information about the process. As shown in Fig. 4(b) and 4(c), the system finds the optimal value by training the process simultaneously during 35 generation. After 35 generation, the error of the network is within the tolerance value, and the optimized value of

the cutting parameters converged steadily. The cutting speed did not change compared to the depth of cut of 0.3mm . However the feed rate was reduced to 0.055mm/rev . This means that main cutting force signal violates the given constraints by increasing the depth of cut. Therefore, the cutting speed should be increased or the feed rate be reduced. However, the spindle speed was constrained by maximum speed of 1200rpm , the feed rate was reduced. This is well coincident to the our intuition.

Fig. 5 shows the surface roughness of the machined work-piece. The surface is very random in the range of initial change of feed rate, on the other hand the surface roughness is constant in the range of 2.

From the above results, we can conclude the system can well learn the process and optimize the cutting parameter regardless of the geometric change such as depth of cut.

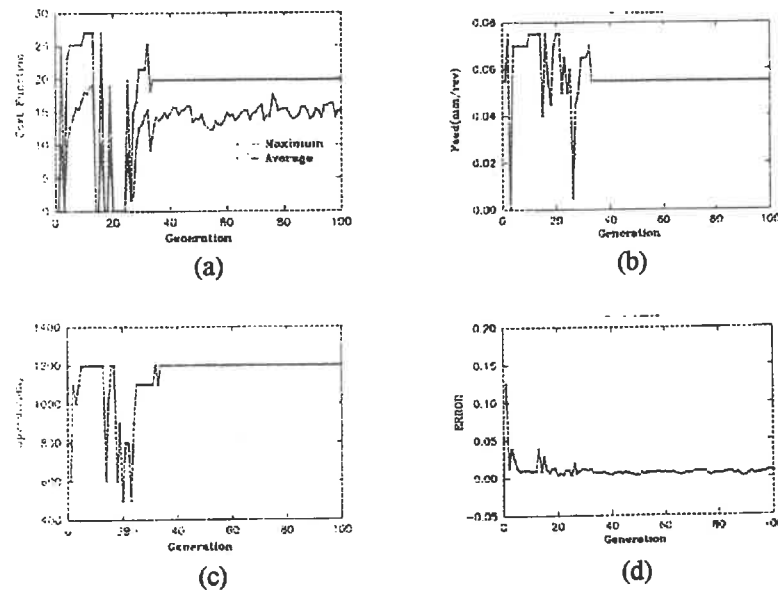
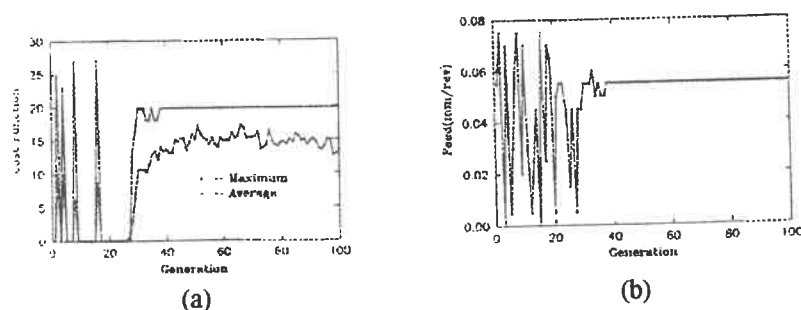


Fig. 4 Results of Genetic Algorithm according to the evolution; (a) Maximum and averaged values of evaluated costs, (b) Variation of feed rate, (c) Variation of cutting speed, (d) Learning errors of a neural network.

Cutting Parameters According to the Progressive Tool Wear

The applicability of the developed system to the progressive change such as tool wear was checked. The experiment was carried out to the new and worn tool. To compare the results, the depth of cut was fixed as 0.3mm .

The maximum and averaged fitness value of the GAs to the worn tool is shown in Fig. 6(a). In this case, the maximum removal rate was obtained as $19.8\text{mm}^3/\text{min}$, which was $7\text{mm}^3/\text{min}$ less than the new tool condition. The network learns the process during initial 30 generation. In this experiment, the tool was artificially ground to fake tool wear. However the optimization may be effective since the network does not learn immediately but progressively. In this case, the feed rate and the spindle speed is 0.055mm/min and 1200rpm , respectively as shown in Fig. 6(b), 6(c).



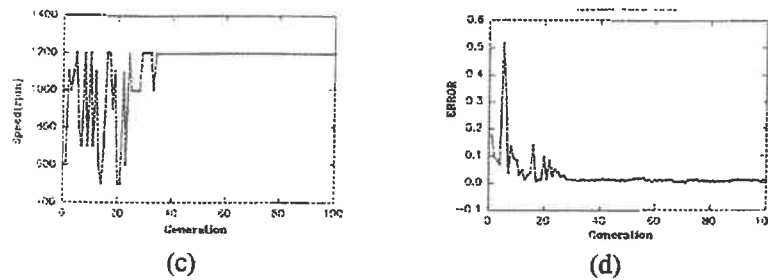


Fig. 5 Results of Genetic Algorithm according to the evolution; (a) Maximum and averaged values of evaluated costs, (b) Variation of feed rate, (c) Variation of cutting speed, (d) Learning errors of a neural network.

CONCLUSIONS

In this research, the adaptive model for the CNC turning lathe was constructed by using an artificial neural network, and the cutting parameters were optimized through the GAs. The system could cut the work-piece optimally regardless of the geometric change such as the change of depth of cut and the progress of tool wear. From the experimental results, we can conclude that the system can well adapt the process change, and find the global optimum cutting parameters without violating the given constraints. The research will be expanded to the surface roughness and cutting temperature to improve the performance.

ACKNOWLEDGEMENT

This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation. The authors would like to thank for the support.

REFERENCES

1. Rangwala, S. S., Dornfeld, D. A., "Learning and Optimization of Machining Operations using Computing Abilities of Neural Networks," IEEE Transactions on SMC, Vol. 19, No. 2, pp. 299-314, 1989.
2. Jang, D. Y., Seireg, A., "Machining Parameter Optimization for Specified Surface Roughness," ASME J. Engng. Ind., Vol. 114, pp. 254-257, 1992.
3. Matsumura, T., Obikawa, T., Usui, E., "Autonomous Turning Operation Planning with Predicting Tool Wear and Surface Roughness," Transactions of NAMRI/SME, Vol. 21, pp. 359-366, 1993.
4. Wu, S. M., Ermer, D. S., "Maximum Profit as the Criterion in the Determination of the Optimum Cutting Conditions," ASME J. Engng. Ind. Vol. 88, pp. 435-442, 1966.
5. Koren, Y., Computer Control of Manufacturing System, McGraw-Hill Book Company, pp. 249-251, 1983.
6. Buckles, B. P., Petry, F. E., Genetic Algorithms, IEEE Computer Society Press, 1992.
7. Michalewicz, Z., Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag, 1994.
8. Goldberg, D. E., Genetic Algorithms in Search, Optimization & Machine Learning, Addison Wesley, 1989.
9. Pao, Y. H., Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Company Inc., 1989.
10. Rao, S. S., Optimization Theory and Applications, 2nd edition, Wiley Eastern Limited, 1984.

Development of a thin brittle filter insertion machine for optical WDM

Kazuki Kudo^{**}, Fumikazu Ohira^{***}, Kunio Koyabu^{***}, Mitsushi Matsunaga^{***}
and Yasuyuki Inoue^{****}

^{**} NTT Multimedia Business Department, Nippon Telegraph and Telephone Corporation,
2-2-2 Otemachi, Chiyoda-ku, Tokyo 100, Japan

NTT Opto-electronics Laboratories, Nippon Telegraph and Telephone Corporation,

^{***} 3-9-11 Midori-cho Musashino-shi, Tokyo 180, Japan

^{****} 162 Shirakata-Shirane Tokai-mura, Naka-gun, Ibaraki 319-11, Japan

Micrometer-order-accuracy insertion of a thin film into a narrow groove using image processing and improved stage accuracy, a method for distinguishing between sides of a curved film, and a precise insertion force control method are proposed and examined in this paper. Using these techniques, an automatic insertion machine has been developed.

To attain a highly-advanced network society that includes Fiber To The Home (FTTH), we will need optical devices such as wavelength division modules that use an optical waveguide or an optical fiber with a thin film filter inserted. Though an optical transceiver incorporating a wavelength division multiplexing (WDM) function is very useful for flexible subscriber access networks, a thin film filter-embedded multiplexer is not still cheap. Thus, an automatic insertion machine that can insert a thin film into a narrow groove reliably and at low cost will be required.

We assume that the filter will be a thin film filter 14- μm thick and that the narrow groove cut into the waveguide will be 18 to 22- μm wide. Such a filter is about 1.4 mm square and consists of a brittle multi-layer and a flexible resin sheet. Consequently, it has a warp of about 10 μm and exhibits both flexibility and brittleness.

Technological problems to be solved are:

- micrometer-order-accuracy insertion of a curved thin film filter into the narrow groove
- handling of the flexible and brittle thin film
- method of distinguishing the sides of the filter
- insertion without damaging the filter

To solve these problems, the following techniques have been newly proposed and examined.

- (i) The relative position of the filter and groove are matched with a high degree of accuracy by

image processing observing from the upper side of the groove. Figure 1 shows a schematic configuration of this insertion method. The waveguide is placed on the 3-dimensional precise stage and its groove is observed from the upper side. Using an appropriate microscopic camera system, the observed image of the groove, as shown in Fig. 2, is clear enough for image processing to determine the position of the groove. This position is matched with the previously-measured position of the filter. The waveguide on the precise stage is transferred to just below the clamped filter. In this case, the repetitive accuracy of transfer movement is crucial. As shown in Fig. 3, the repetitive accuracy of the transfer movement was highly improved to about $1\mu\text{m}$.

(ii) The film is sucked by the vacuum clamp or clamped mechanically to correct the warp.

(iii) To distinguish the sides of the film, the film surfaces are scanned by a laser beam to measure their shape. Examples of scanned film profiles are shown in Fig.4, indicating that the sides of a filter can easily be distinguished.

(iv) To avoid damaging the filter during insertion, a weight cancellation function for the insertion mechanism by electromagnetic force is applied to control insertion force precisely. To optimize insertion force, the minimum force for insertion and the maximum force, above which the filter is damaged, are measured. The minimum force is 3 gf and the maximum force is 10 gf.

With all the techniques proposed above, a precise and reliable insertion is achieved. Moreover, an automated machine based on these techniques has been made. (Fig. 5) The machine picks up parts with its hand unit from the parts tray, inserts them and bonds them in place using adhesive and UV curing. UV adhesive is supplied by capillary action.

Using this machine, tolerance of the relative positioning error between centers of the filter and the groove for successful insertion is measured. The measured tolerance is about $25\mu\text{m}$. This width is several microns wider than that of the groove. This means that insertion is successful even when the filter touches the edge of groove. The average time for this machine to produce a module is about 1 minute. The module's light loss for the $1.5\mu\text{m}$ wavelength reflected from the inserted filter is around 1.5 dB, which satisfies the requirements of less than 2 dB.

This machine should be capable of producing a large number of indispensable optical devices for the highly-advanced information society of the future at low cost.

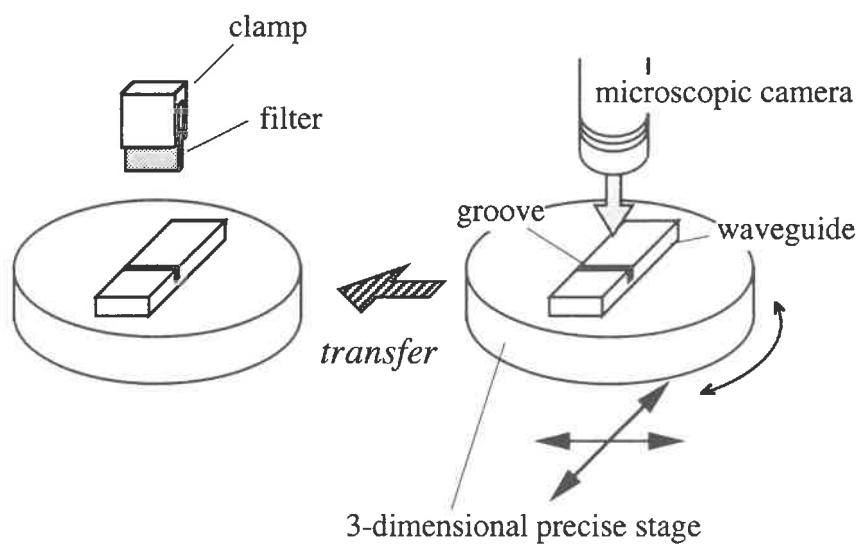


Fig. 1 schematic configuration of insertion method

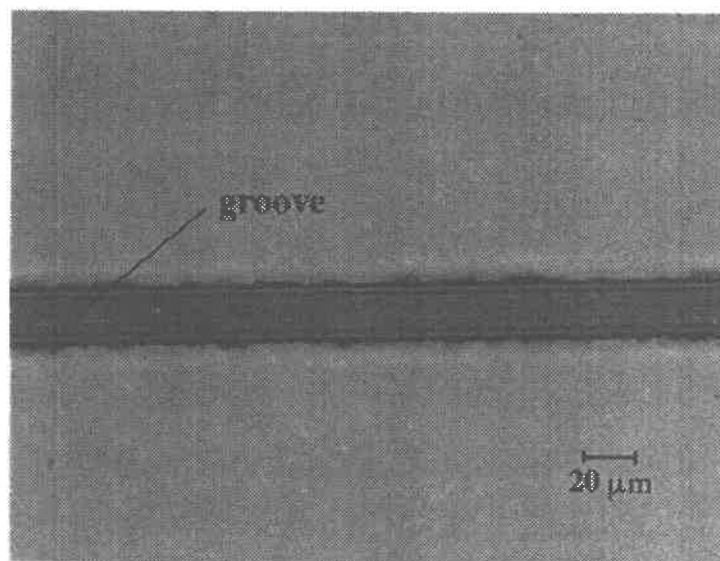


Fig. 2 image of a groove

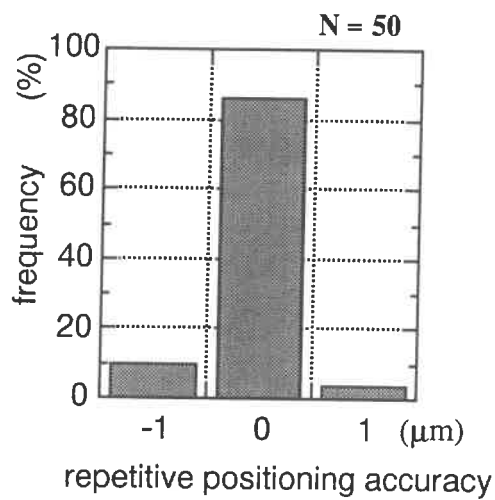
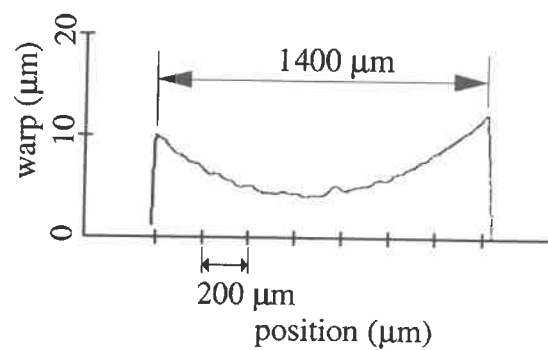
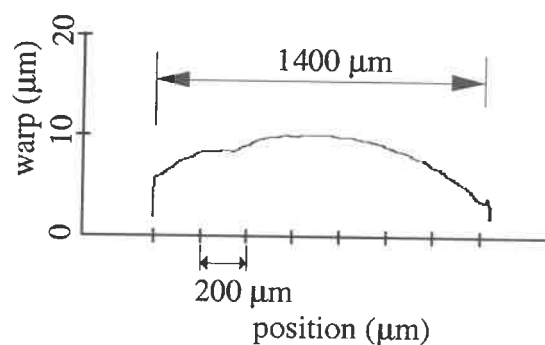


Fig. 3 repetitive positioning accuracy of transfer stage



(a) sample #1



(b) sample #2

Fig. 4 scanned film profiles

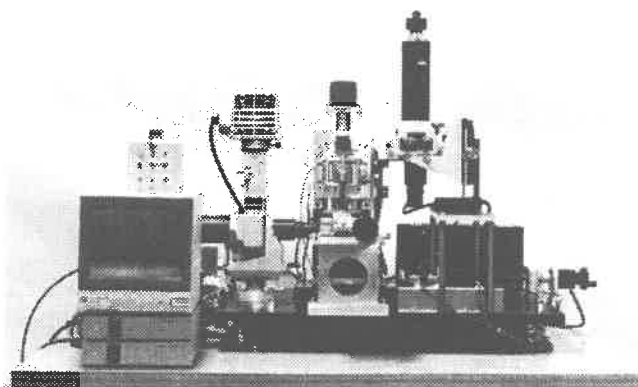


Fig. 5 photograph of developed filter insertion machine (prototype)

Autonomous acquisition of a shared, machining condition data-base in an open CNC environment

Mamoru MITSUISHI, Hideyuki OKABE, Makoto HASHIGUCHI, Katsuya TANAKA
and Takaaki NAGAO

Department of Engineering Synthesis, Faculty of Engineering
The University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan

1. Introduction

The authors have developed special functions that enable highly efficient adaptive control for high precision cutting with high cutting stability[1]. The system uses multi-axis force information that can be incorporated with control results and autonomously accumulated as a data-base in the form of a stability lobe diagram. The data-base is renewed in real-time and allows the optimization of the cutting conditions, including tool path, quasi-dynamically. Conventionally, each individual machine tool has its own data-base because the stability of the system depends on the stiffness and the damping factor of the machine. However, it is more efficient if the machining parameter data-base can be shared among multiple machines. This paper discusses a common data-base renewal method which considers the effects of the differences in mechanical properties among machine tools and of tool wear.

2. Related work

Many researchers have discussed the mechanism of chatter vibration and developed physical models[2, 3, 4]. In particular, Sato, et al. are discussing the multiple regenerative effect of chatter vibration[5]. [6] and [7] discussed a monitoring method of tool wear and breakage based on acoustic emission signals. A method for machining state detection and the modification of the conditions by adjusting the tool length was presented by Smith and Thusty[8]. Some researchers are discussing open-architecture CNC machine tools [9, 10, 11]. In particular, Altintas, et al. realized an adaptive control on a machine tool equipped with an open-architecture CNC[10]. However, in his system the machine tool is not closely connected with CAD/CAM system. In our system CAD/CAM system and the machine tool are tightly connected[1]. In the system the optimization of cutting conditions is performed by changing the data-base dynamically during cutting. This paper describes a

sharing method of data-base modification strategies depending on tool wear between two different machining centers.

3. System construction

3.1 Hardware system

A machining center which is controllable from an external, real-time computer, developed by the authors' group, was used. The machining center is equipped with a 6-axis force sensor to monitor the cutting state and a fail-safe table to prevent damage to the system in the case, for example, of an over-force condition. The rigidity, the natural frequency and the rated value of the 6-axis force sensing table developed by the authors' group are approximately 500N/ μ m, 2kHz and 2,000N for the spindle rotation direction and 1,000N for the orthogonal directions, respectively. The real-time controller, the data-base and the CAD system are connected and integrated with each other through a computer network.

3.2 Software system

Fig.1 shows an overview of the software system. The system consists of a CAD system, a data-base and a real-time controller. One of the features of the system is that these processes are not static, but dynamic, communicating with each other during the cutting sequence. More concretely, the contents of the data-base are dynamically revised by being sent the actual cutting conditions and cutting state from the real-time controller. Furthermore, the three-dimensional modeller in the CAD system can be accessed dynamically during cutting. Tool path and cutting conditions are regenerated depending on the revision of data-base and are sent to the real-time controller.

3.2.1 Data-base

The data-base consists of an "experimentally obtained stability lobe diagram," a "machining efficiency evaluation task" and a "machining condition determination task." The stability lobe diagram introduced by Merritt[2] and Thusty is

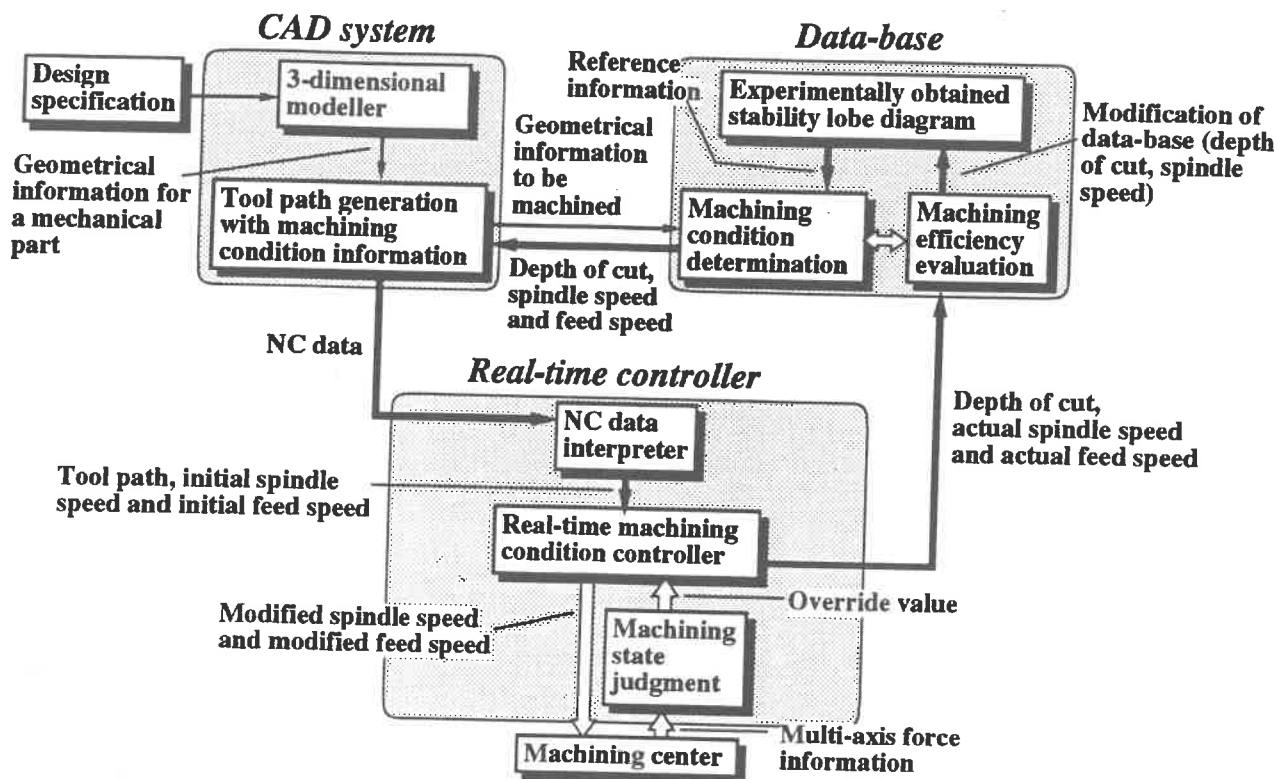


Fig.1 Software system configuration

experimentally obtained in the system described in this paper. Therefore, it is called an "experimentally obtained stability lobe diagram." Fig.2 shows an example of a stability lobe diagram.

The machining efficiency evaluation task and the machining condition determination task use the 3-dimensional modeller in the CAD system to determine efficient machining conditions and generate an appropriate tool path. Furthermore, the machining efficiency evaluation task modifies the stability lobe diagram as it varies during ma-

chining.

3.2.2 Real-time control

The "real-time machining condition controller" and the "real-time machining state judgment task" perform adaptive control by adjusting the override values for spindle speed and feed speed based on the previously determined machining conditions. Furthermore, the relationship between the machining conditions and the machining state is fed back to the data-base.

3.2.3 CAD system

In the CAD system, geometrical information for the component mechanical parts of a product is input to the 3-D modeller and the abstracted machining feature information is sent to the machining condition determination task in the data-base. The tool path generation task generates a 3-dimensional tool path appropriate to the received depth of cut and geometrical machining feature information. The most efficient set of spindle speed and depth of cut obtained from the machining condition determination task changes when the stability lobe diagram in the data-base is modified. Therefore, in the system described in this paper, the tool path generation task is invoked periodically during machining and the tool path and cutting conditions are quasi-dynamically modified to maintain maximum cutting efficiency. Typically, the tool path is regenerated after each cutting pass.

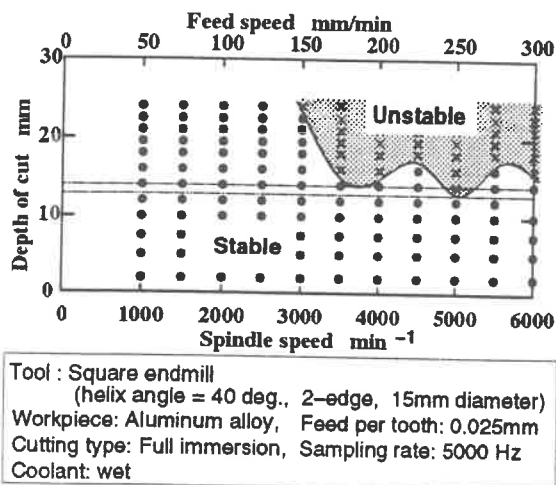


Fig.2 Stability lobe diagram

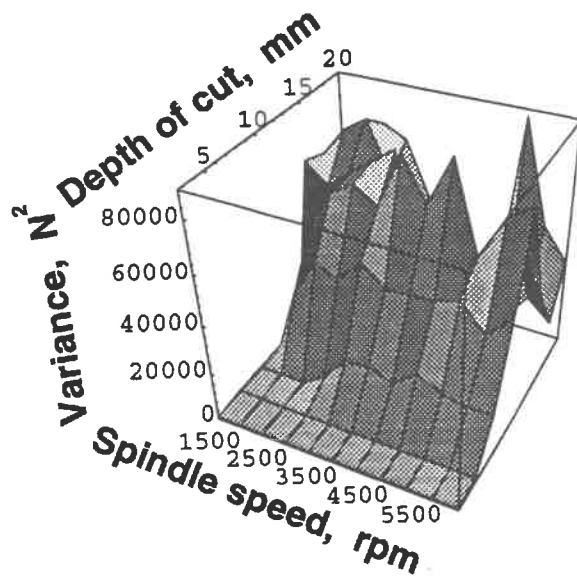


Fig.3 Variance of distance for M.C.-A (sharp tool)

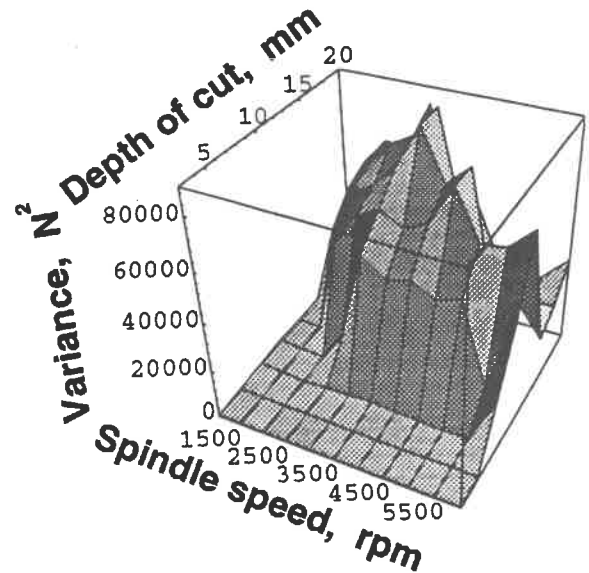


Fig.4 Variance of distance for M.C.-A (significant tool wear)

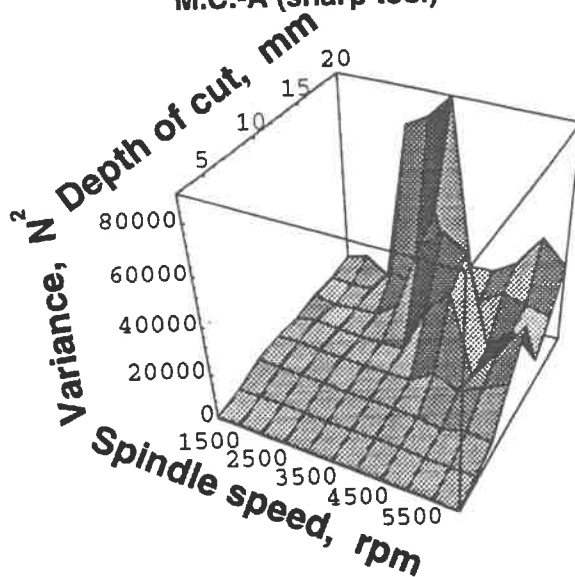


Fig.5 Variance of distance for M.C.-B (sharp tool)

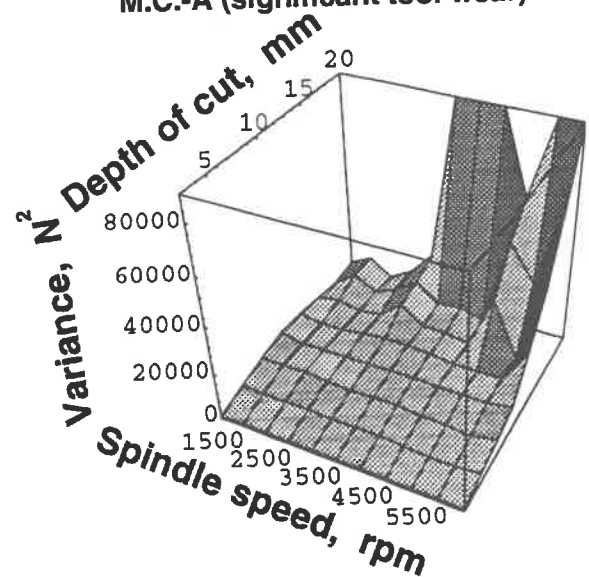


Fig.6 Variance of distance for M.C.-B (significant tool wear)

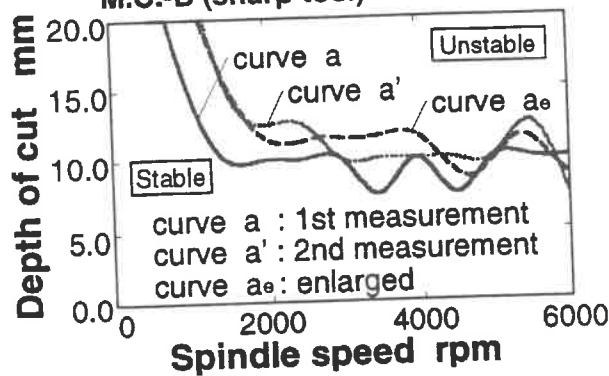


Fig.7 Modification of stability lobe diagram for M.C.-A

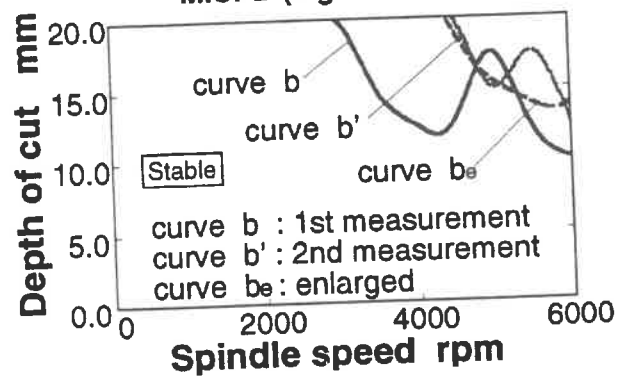


Fig.8 Modification of stability lobe diagram for M.C.-B

Table 1 Cutting conditions

Tool	square endmill (10 mm diameter, helix angle 40 deg., 2 teeth)
Workpiece	Aluminum alloy(A2017)
Feed per tooth	0.025mm/tooth
Cutting type	Full immersion
Sampling rate	5kHz
Coolant	Wet

4. Data-Base sharing

As cutting stability of a system depends on mechanical properties of machine tools such as shape of cutting tools, geometical shape of workpieces and machine tool structures. The stability lobe diagram is different for each machine tool. Therefore, it is required to obtain diagrams separately for various combinations of factors mentioned above. However, it is inefficient. One solution to solve this problem is to share data-base among multiple machine tools. According to [3], parameters which depend on tool wear are chip-thickness coefficient k_1 and penetration-rate coefficient K . k_1 transforms the diagram in the direction of depth of cut. Furthermore, it is known that K enlarges or shrinks the lobe in both the spindle speed and depth of cut directions. Therefore, the diagram after significant tool wear will be obtained by enlarging and shrinking the original one in the respective directions. Data-base renewal method is shared by considering that the magnification ratio presents the amount of the tool wear.

5. Experiment

Sharing of data-base renewal method between two machine tools was tried in the experiment. At first, the transition of stability lobe diagram caused by tool wear was measured on one machine tool. The tendency of lobe changes was used to estimate the other machine's lobe change after significant tool wear.

Fig.3 and Fig.4 were obtained on machine tool A. Cutting conditions in the experiments are shown in Table1. The bases of the figures are composed of spindle speed and axial depth of cut, and the vertical axis represents a variance of distance in force vector locus which indicates the machining state. The cutting state is stable if variance of distance is small. Fig.3 and Fig.4 were obtained with a sharp tool and the same tool after significant tool wear, respectively. Fig.5 and Fig.6 show experimental results on machine tool B with the same condition as Fig.3 and Fig.4. Fig.7 shows the cross sections of Fig.3 and Fig.4 at the height of 20,000N². Curve a_e was selected so as to minimize the area enclosed by curve a' and a curve obtained

by magnifying curve a . In this experiment, the magnification ratio were 1.35 and 1.15 in the directions of spindle speed and axial depth of cut, respectively. Curve b and curve b' in Fig.8 are the stable-unstable boundaries in the cross sections of Fig.5 and Fig.6 at the height of 20,000N². Curve b_e was obtained by magnifying curve b at the same ratio as the case of Fig.7. Curve b_e can be regarded as the prediction of curve b . As is shown in Fig.7, stability lobe is well predicted. In particular, in the range of lower than 5,120 rpm, the predicted curve fits exactly to the measured one.

6. Conclusions

The authors proposed a method for sharing machining data-bases among different machining centers whose rigidities and damping factors are different from one another. In the experiment, the proposed method was applied for the effect of tool wear. The experimental result shows the effectiveness of the proposed method.

References

- [1] Mitsuishi, M., et al., "A Practical Machining Condition Determination Strategy Using Multi-Axis Force Information," *Annals of the CIRP*, 45/1, pp.373-376, 1996.
- [2] Merritt, H.E., "Theory of Self-Excited Machine Tool Chatter," *ASME J. Eng. for Industry*, pp.447-454, Nov. 1965.
- [3] Tobias, S.A., "Schwingungen an Werkzeugmaschinen," *Carl Hanser Verlag*, München, 1961.
- [4] Minis, L., et al., "Analysis of Linear and Non-linear Chatter in Milling," *Annals of the CIRP*, 39/1, pp.459-462, 1990.
- [5] Kondo, Y., Kawano, O. and Sato, H., "Behavior of Self-Excited Chatter due to Multiple Regenerative Effect," *ASME, J. Eng. for Industry*, 103, 3, pp.324-329, Aug. 1981.
- [6] Moriwaki, T., et al., "Application of Neural Network to AE Signal Processing for Automatic Detection of Cutting Tool Life," *Journal of JSPE*, vol.57, no.7, pp.1259-1264, 1991.
- [7] Dornfeld, D.A. and Liu, J.J.B., "Abrasive Texturing and Burnishing Process Monitoring Using Acoustic Emission," *Annals of the CIRP*, 42/1, pp.397-400, 1993.
- [8] Tlustý, J., et al., "Techniques for the Use of Long Slender End Mills in High-speed Milling," *Annals of the CIRP*, 45/1, pp.393-396, 1996.
- [9] Pritschow, G., et al., "Open System Controllers - A Challenge for the Future of the Machine Tool Industry," *Annals of the CIRP*, 42/1, pp.449-452, 1993.
- [10] Altintas, Y., et al., "A Hierarchical Open-Architecture CNC System for Machine Tools," *Annals of the CIRP*, 43/1, pp.349-354, 1994.
- [11] Koren, Y., et al., "Real-Time Open Control Architectures and System Performance," *Annals of the CIRP*, 45/1, pp.377-380, 1996.

REVIEW OF DEVELOPMENTS IN SENSORS FOR UNTENDED MACHINING

A. Chukwujekwu Okafor and Anietie Ukpung

Laboratory for Industrial Automation and Flexible Manufacturing

Department of Mechanical and Aerospace Engineering and Engineering Mechanics

University of Missouri-Rolla

Rolla, MO 65409-0050

ABSTRACT

Traditional manual quality control based on detection after the facts — a post mortem process — are no longer adequate in the wake of increasing demand on manufacturing productivity forced by domestic and global competition. Costly and abrupt failures resulting from undetected tool wear, collision or breakage should be a thing of the past and moreover, the pressure of cost is increasingly forcing the service life of manufacturing equipment to be fully utilized, and tools are not to be changed or reground prematurely simply on the basis of experience or for safety reasons. Since the early 80s, when a comprehensive review of manufacturing sensors for unmanned machining was carried out by J. Tlustý and G. C. Andrews, numerous developments in sensors for unmanned machining have taken place. This review highlights these developments — past and present. The review appraises sensors for machining monitoring on the basis of specification and characteristics, reliability, precision, signal processing, speed and limitations. The extent to which these sensors are effective individually or in combination as a machining monitoring system are discussed.

1. INTRODUCTION

It has been established, that in order to increase the productivity of computer numerical control (CNC) machines, untended or minimally supervised operating conditions are desirable [1]. The emergence of unmanned machining centers and problems associated with adaptive control or untended machining in an operator's absence has continued to require attention. An untended machining situation thus warrants that operations provide for artificial intelligence necessary to mimic the qualitative judgement of operators and their instinctive reactions and, provide the expertise normally associated with human involvement. There have been numer-

ous reviews on machining sensors and machine tool monitoring systems [1-3]. There has been however, a passive attitude towards the integration or incorporation of these technologies into our manufacturing endeavours, such that we are still currently faced with a situation of *backlog of unapplied technologies*. Moreover, there is still the reluctance of machine tool builders for changes in machine tool design configurations to incorporate the various "eyes" and "fingers" called sensors requisite for untended manufacturing. The review of these sensors are given in the following sections.

2.0 Types of commercially available sensors for untended machining

All the prerequisites for untended machining calls for operations such as automatic parts loading and unloading, in-process part gauging, adaptive control, monitoring for tool wear, collision, and breakage detection. Direct measurement techniques have been established to pose many problems for automation. Efforts therefore are directed toward indirect measurements of process variables that are closely correlated with tool and machining conditions.

2.1 Touch trigger probes

This device is used for workpiece monitoring and the information which is provided may be used to update tool offsets critical in an untended machining environment. First patented by Renishaw, the system consists of a very sensitive omnidirectional switch made up of a probe body and a stylus capable of detecting deflections in any direction. The probe body is available in two styles enabling machining center spindle-mounting and, machine tool table mounting for milling and drilling (figure 1.) For turning centers, the probe is mounted on the turret and programmed to confirm that the correct workpiece is loaded. Using results of probing and conditional part programming,

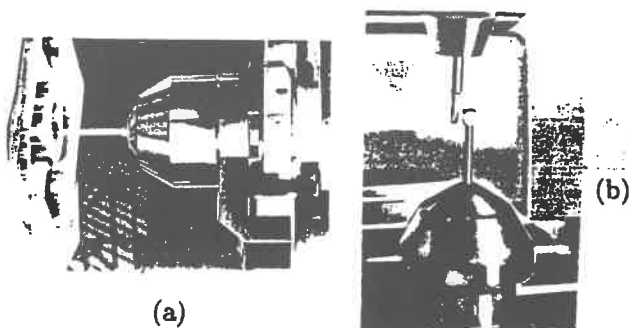


Figure 1: (a) spindle-mounted probe for machining center (b) tool-table mounted probe for drilling and milling

various workpieces can be accommodated, identified, and defined. Typical operations for CNC machine tools equipped with touch-trigger probes are stated to include: tool setting, initial job set-up, and sensing of tool wear and breakage. The probe can be used for nearly all tasks demanded for unmanned machining, however, the question of cost and efficiency, response time, and memory becomes critical and limits them to certain metal cutting operations [2].

2.2 Cutting force and torque sensors

There are many ways of measuring and inferring cutting forces. However, most of the techniques are still limited to laboratories or their implementations are ad hoc. Commercially developed systems for unmanned machining over the years have tended to be force/torque - based or acoustic emission - based. Ways of measuring or inferring cutting forces or torques for unmanned machining include:

- tool-holder and table-type dynamometer
- rotating cutting force dynamometer
- dynamometer built into spindle bearings
- 3-component piezoelectric measuring pin
- piezoelectric strain transducers
- force measuring bushings and bearings
- inferring cutting force from spindle torque, electric current and power consumption of the main or feed spindle motors.

2.2.1 Dynamometers

Recent developments in piezoelectricity have produced numerous sensors with potential for use in tool monitoring. Kistler and Montronix are among the industrial leaders utilizing this technology in the development of a wide range of sensors by exploiting the piezoelectric effects that can be activated in quartz crystals.

(a) Tool holder dynamometers

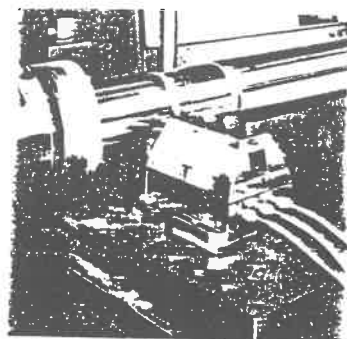


Figure 2: tool holder dynamometer

Figure 2 shows a tool holder dynamometer for typical force measurement with a 3-component dynamometer. As observed [2], these types of sensors are suitable for several tasks because of high sensitivity and high natural frequency, but they can not be used in turning centers with turret heads. Furthermore, being a specific laboratory tool, its bandwidth is still not large enough for high frequency force variation due to chip formation especially during high speed cutting. Other tool holder dynamometer based on strain gauge with small ring sensing element have been developed by WZL Aachen [2].

(b) Three component measuring pin

To circumvent problems associated with inadaptability of tool holder dynamometers for turret lathes, Kistler and Montronix have introduced a new 3-component measuring pin. It consists of a piezoelectric 3-component force measuring cell, and a double preloading wedge, figure 3 [5], enabling indirect measurement of three-force components orthogonal to each other. In principle, a three-component measuring pin of this kind measures the deformation caused in machine parts by external forces and torques. A typical Kistler's 9117A measuring cell is also shown in figure 3. The basic requirement for using this sensor is to ensure that the transducer is installed in the load path of the machine tool and in as close proximity as possible to the point of cut.

(c) Table - type dynamometers

Most of these are available as a measuring platform, which have an advantage that they permit a larger freedom of point of load application than would a single transducer [4]. Measuring platform 9257A for a 3-component force transducer from Kistler is shown in figure 4. On lathes, the platform is mounted on the slide and the tool is bolted onto the platform, whereas on milling machines, the workpiece is bolted onto the platform which is then affixed onto the machine table. It is conceivable that table - type dynamometers will

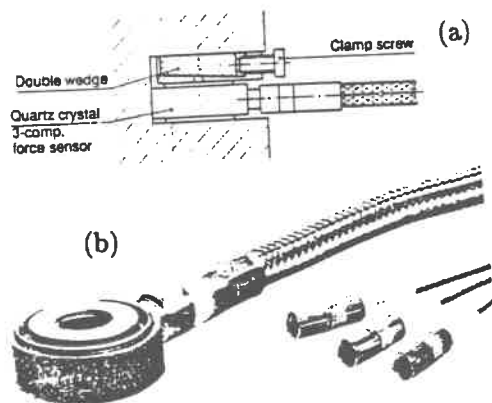


Figure 3: (a) measuring cell with preloading wedge, (b) Kistler's 9117A

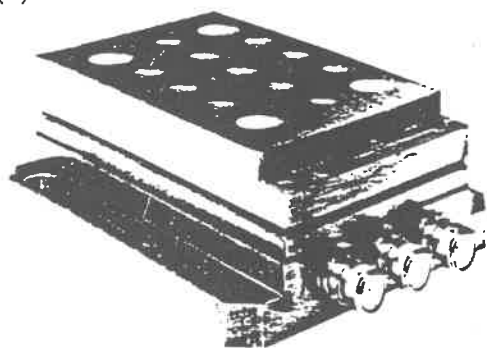


Figure 4: Kistler's 9257A measuring platform

be used for adaptive control in milling and for providing signals for tool wear and breakage detection simultaneously. However, for chatter research and testing of machining centers, special wide frequency band dynamometers are needed [2].

(d) Rotating cutting force dynamometers

The new rotating cutting force dynamometer (RCD) introduced by Kistler has the capability of measuring forces and moments on the rotating cutting edge, figure 5. Its high rigidity and compact design makes it adaptable to all machine tools and a large assortment of tools. The forces occurring (F_x , F_y , F_z) and moment (M_z) are measured with a piezoelectric sensor, featuring high rigidity, wide measuring range, and a minimum response threshold [6]. The measuring system has a facility for amplifying and transmitting the signal to a stationary receiver using pulse modulated telemetry. In milling, the RCD can be used for both exact precision and coarse processes by simply switching over measuring ranges. A force measuring system to determine tool-related cutting forces in 5-axis milling is shown in figure 6 [7]. The flange design of the tool holder enables tools most frequently used to be accommodated.

(e) Dynamometers built into spindle bearings

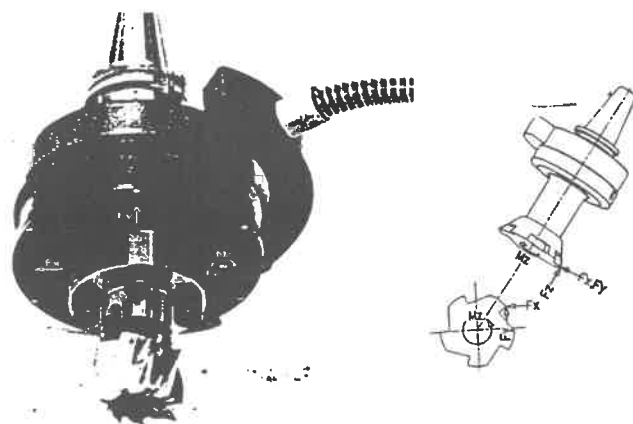


Figure 5: RCD and RCD coordinate system

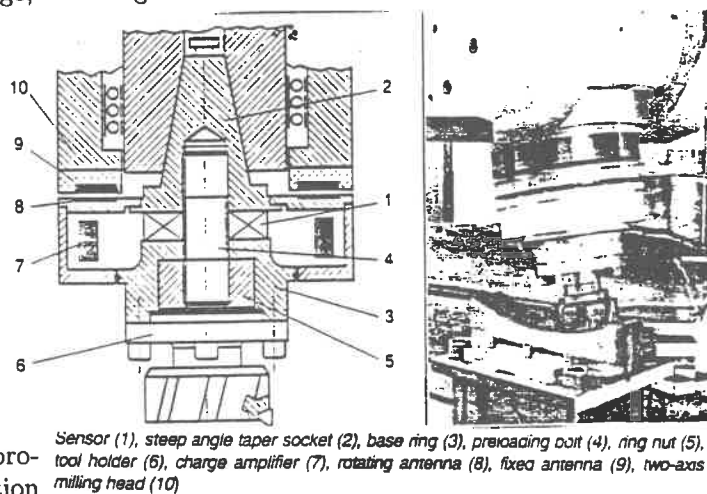


Figure 6: RCD system for 5-axis milling

The first break through by PROMESS was to have spindle bearings communicate load information [8]. As shown in figure 7, when a roller bearing (tapered or ball) is loaded, load proportional strains are taken up by both inner and outer races of the bearing, leaving a major portion of the forces being transmitted in the hertz area. As load varies, this AC signal varies as a result of deformation variation to produce amplitude modulation, definable as axial or radial forces by strategic placement of the sensors. It is this change in amplitude that measures the forces. A major drawback of force measuring bearing, is the need to eliminate rolling noise frequencies by using a low pass filter, which consequently makes it impossible to detect rapid signal changes [3]. Further innovations for the measurement of fast signals include force-measuring bushings figure 7, and the fact that they require changes in the design of bearing arrangement make retrofitting almost always impossible. A major application area however, is in the axial force measurement on drilling machines. The cross-section of a typical assembly of the Z-axis feed force sensor built into spindle bearing to measure thrust force

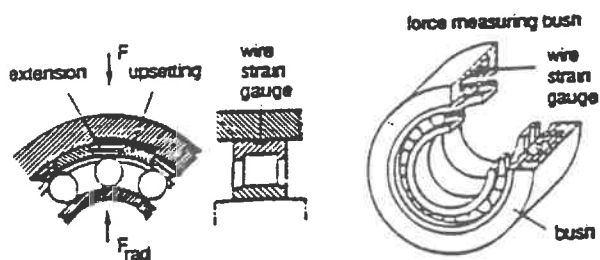


Figure 7: dynamometer built into spindle bearings and force measuring bushing

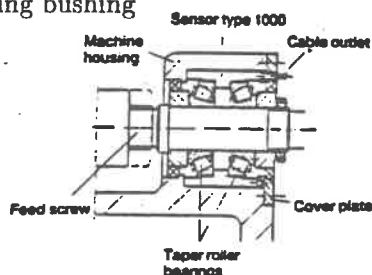


Figure 8: Z-axis feed force sensor

from Sandvik is shown in figure 8. This sensor can be used for the mounting of thrust bearings of a lead screw to monitor feed forces and to detect drill wear or any other effect which alters feed force, such as deterioration of the guideways [2]. The sensors are available in sizes from 25mm to 350mm and from 2 kN to 100 kN force ranges. Their compact design for machine tools ensures easier installation.

2.2.2 Inferring cutting force from spindle torque, electric current and power consumption of the main or feed spindle motors

These indirect ways of measuring cutting forces and torques have been used in the development of many commercial applications, including the well documented Cincinnati Milacron's torque control monitor (TCM) [2,9]. It is easy to measure the motor current and voltage and thus evaluate the power being used and, by sensing the motor speed with current and voltage, it is possible to evaluate torque. Hence, with the diameter of the workpiece during turning or of the cutter in milling, the tangential component of the cutting force can be determined [2]. The problem with this approach is that the power required to overcome friction in the guideways usually far exceeds that needed for machining and the output signal usually has a low pass filtering characteristic due to inert masses. Moreover, tool wear monitoring is often difficult as spindle power is proportional to the resulting cutting force in the direction of primary motion, the least tool wear sensitive parameter. Despite these problems, such devices and techniques are employed

commercially to detect tool collision, breakage, and wear if adequately large cross-sectional areas of cut are used [2].

2.3 Acoustic emission sensors

During metal cutting operations, acoustic emission (AE) results from the plastomechanical processes of crack formation, growth, chip removal, and from surface friction in the cut zone [3]. The exploitation of the AE signal for tool wear and breakage detection research is intensive [11]. Because of the propagation of waves through the whole machine structure, the point of attachment of an AE sensor is not very critical. However, multiple joints between the cutting position and the AE sensor result in increased attenuation of the AE signal on the order of 10db per interface, thus the sensor must be located as close as possible to cutting action [3]. Several AE sensors have been developed (figure 9). Digital Wave Corp. has introduced a new AE analysis technique based on waveform analysis using broadband transducer to capture wide range of frequencies within a given wave. The technique has unique merits over conventional AE analysis technique using resonant sensors [18].

A Piezotron AE sensor developed by Kistler, figure 10, has a built-in impedance converter for measuring AE above 50 KHz in machine structures. With its small size, it mounts easily near the source of emission and optimally captures the signal [12]. The power spectra of AE signal during normal cutting and at fracture is shown in figure 11 and it is obvious that the tool fracture signal releases high-powered oscillations in frequency range 300 KHz to 1 MHz, and it is distinguished from normal cutting bursts if passed through a high-pass filter with 300 KHz corner frequency. All other low frequency signals, such as those due to audible noise and mechanical vibrations would be filtered out at the same time. This is the basis of tool breakage detection by AE [2].

2.4 Dual-mode force/AE sensor

Recently, Kistler has introduced a dual-mode sensor that permits the simultaneous measurement of AE signals and one to three orthogonal forces, figure 12 [10]. Integrated in the middle of a force measuring ring, is an AE sensor with same properties as those of a Piezotron. The cutting force and AE signal recorded simultaneously during external diameter lathing is shown in figure 13. A dual-mode sensor was mounted under the turret housing and tool tip breakage signal with that recorded during turning over a hole are shown. The signals change due to interrupted cuts are easily recognized in the force signal

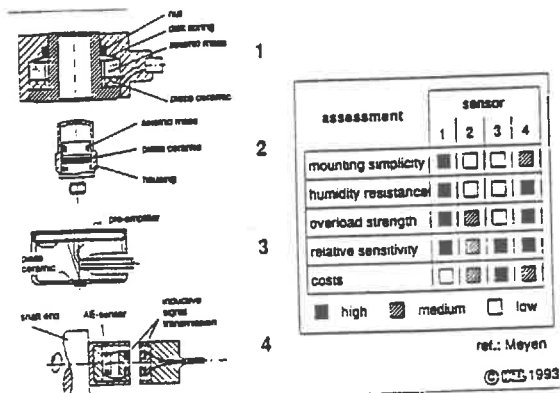


Figure 9: AE Sensors

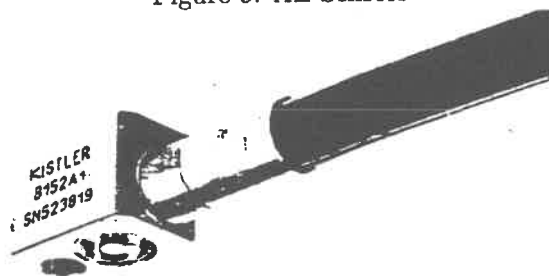


Figure 10: Kistler's Piezotron AE sensor

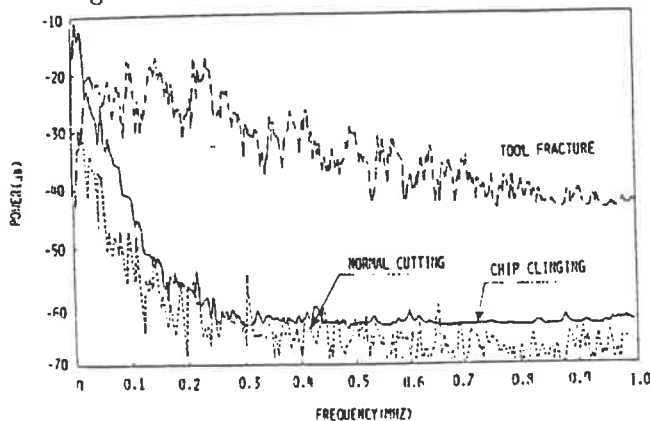


Figure 11: power spectra of AE signal

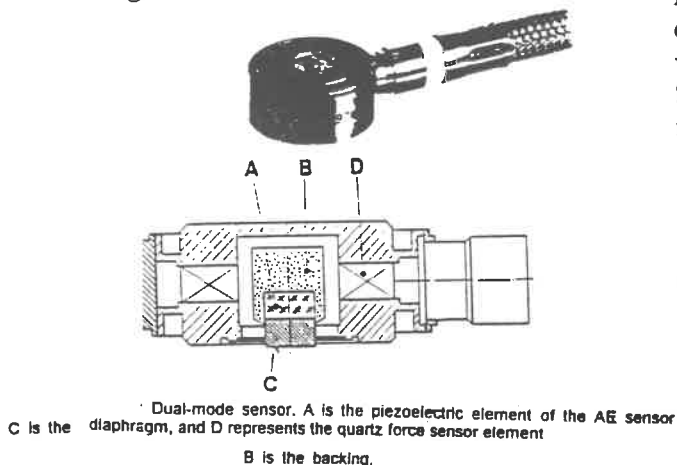


Figure 12: dual-mode AE sensor

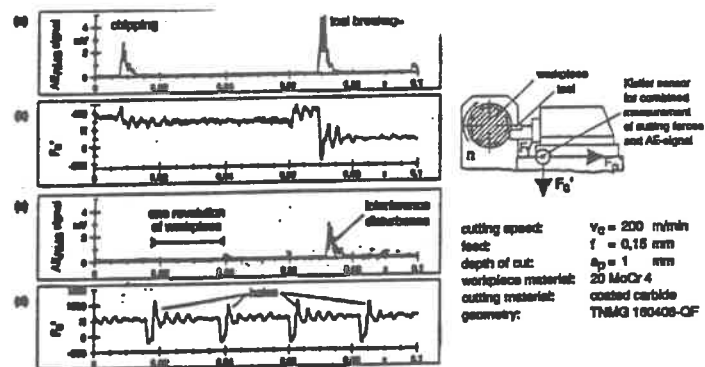


Figure 13: cutting force and AE signal recorded simultaneously

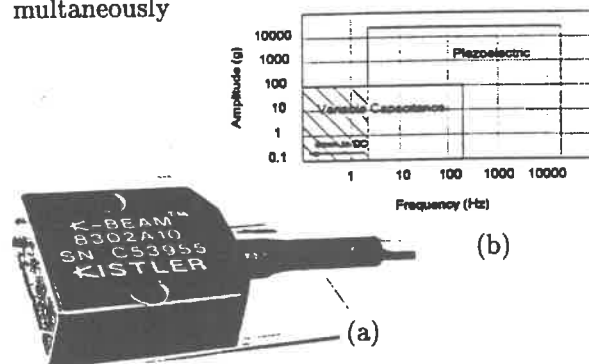


Figure 14: (a) K-BEAM variable capacitance accelerometer (b) Dynamic range comparison

but not in the AE. However, the pronounced peak in the AE signal caused by an interference is also not shown in the force signal. This example shows how multi-sensor signals allows a more reliable detection of the actual process disturbance – tool breakage.

2.5 Accelerometers

Kistler has recently introduced the K-BEAM silicon micromachined capacitive accelerometer for measuring low frequency, low level acceleration to zero Hertz. Figure 14 shows the variable capacitance accelerometer along with its dynamic range comparison with a piezo-electric type. They are available in $\pm 2g$, $\pm 10g$ and $\pm 20g$ ranges with high signal-to-noise ratio. Both types have unique merits [17].

3.0 SUMMARY

Many sensors and techniques have been studied and developed for tool wear or failure detection over the years. Most of these sensors and techniques can not by themselves suffice to provide the much needed unmanned machining and adaptive control edge. Different sensors are often sensitive to different process changes [14]. For instance, AE is less sensitive to cutting conditions such as depth of cut and feed, while force is less sensitive at reduced chip cross-sectional

area [3]. Hence the need for multiple sensors to take advantage of differing sensitivities of sensors to increase the range of applicability of sensing system. Pioneering efforts in this direction are well established [13-15]. Open issues must include not only the development of new sensors, but also the ability to use existing sensors in better ways to effectively understand the physics of chip formation [16] and tool conditions. We must also be able to combine outputs from such sensors into software architectures that more closely model the human machinist's ability to decide what to do next given the current state of the manufacturing process – A motivation for neuro-fuzzy tool monitoring systems.

4.0 ACKNOWLEDEMENT

The financial support of this work by the National Science Foundation, NSF grant No. GER-9355127 is gratefully acknowledged.

5.0 REFERENCES

1. Smith, Graham T., Advanced Machining – The Handbook of Cutting Technology, IFS Pub., pp. 80-107, 1989.
2. Tlustý, J. and Andrews, G. C., "A Critical Review of Sensors for Unmanned Machining", *Annals of the CIRP*, Vol. 32, pp. 563-572, 1983.
3. Dornfeld, D. A., König, W. and Ketteler, G., "Current State of Tool and Process Monitoring in Cutting, New Developments in Cutting Technology," *Proc. of CIRP/VDI Conf.*, Dusseldorf, Vol. 998, pp. 363-376, 1993.
4. Gautschi, G. H., "Cutting Forces in Machining and Their Routine Measurement with Multi-Component Piezo-electric Force Transducers," *Proc. of the 12th Int. Mach. Tool Design and Research Conf.*, Manchester, England, pp. 113-120, Sept. 1971.
5. Cavalloni, C., Engeler, P. and Schaffner, G., "Strain Measurement Inside Structures with Piezoelectric Sensors," *Sensors and Sensor Systems*, Expert Verlag, D-7044, Ehnningen/Germany, pp. 132-143.
6. Cselle, T. and Stirnimann, J., "Cutting Force Measurements on Rotating Tools," Translated Paper, Carl Hanser Verlag, Munchen, 1993.
7. Spur, G., Al-Badrawy, S. J. and Stirnimann, J., "Measuring the Cutting Force in Five-Axis Milling," Translated paper, Carl Hanser, Munchen, Courtesy of Kistler Piezo-instrumentation, Sept. 1993.
8. Stockline, Larry E., "Tool and Machine monitoring Systems," *4th Biennial Int. Mach. Tool Tech. Conf.*, Vol. 10, pp. 158-165, Sept. 1988.
9. Burgam, P. M., "Two Tools for the Eighties: Machining Centers and Adaptive Control," *Manufacturing Engineering*, pp. 56-60, May 1984.
10. Claudio, Cavalloni and Kirchheim, Andreas, "New Acoustic Emission Sensors for In-Process Monitoring," *Presented at the 12th. Int. Acoustic Emission Symposium*, Sapporo, Japan, Oct. 17-20, 1994.
11. Dornfeld, D., "Acoustic Emission and Metal-Working Survey of Potential and Examples of Applications," *Proc. of the 8th NAMRC*, University of Missouri, Rolla, May 18-21, Dearborn, Michigan, pp. 207-213, 1980.
12. Kistler Instrument Corporation, "Piezotron," In-House Print, 1994.
13. Dornfeld, D. A., "Neural Network Sensor Fusion for Tool Condition Monitoring," *Annals of the CIRP*, Vol. 39, pp. 101-270, 1990.
14. Okafor, A. C., Marcus, M. and Tirpirneni, R., "Multiple Sensor Integration Via Neural Network for Estimating Surface Roughness and Bore Tolerance in Circular End Milling - Part 1: Time Domain," *Condition Monitoring and Diagnostic Tech. Journal*, Vol. 2, No. 2, pp. 49-57, 1991.
15. Mesina, O. S. and Langari, Reza, "A Neuro-Fuzzy System for Tool Condition Monitoring in Metal Cutting," *ASME Dynamic Systems and Control*, Vol. 2, pp. 931-938, 1994.
16. Wright, P. K. and Bourne, D. A., *Manufacturing Intelligence*, Addison-Wesley Pub. Co. Inc., 1988.
17. Berther, T., Kubler, J. M. and Cuhat, D., "A Lightweight Capacitive Accelerometer for Low Frequency Measurement," Kistler Piezo-Instrumentation, In-House Print, 1994.
18. Okafor, A. C., Chandrashekhara, k. and Jiang, Y. P., "Location of Impact in Composite Plates Using Waveform-based Acoustic Emission and Gaussian Cross-correlation Techniques," in *Smart Structures and Materials 1996: Smart Sensing, Processing and Instrumentation*, Kent A. Murphy, Dryver, R. Huston, Eds., Proc. SPIE 2718, pp. 291-302.

Integrated Electrodynamic Multicoordinate Drives - Modern Components for Intelligent Motions

Dipl.-Ing. Chr. Schäffel; Dipl.-Ing. E. Saffert; Univ.-Prof.Dr.-Ing. habil. E. Kallenbach;
Technical University of Ilmenau, Faculty of Mechanical Engineering, Germany
e-mail: christoph.schaeffel@maschinenbau.tu-ilmenau.de

ABSTRACT

In the next decades the machine precision will permanently increase in both standard and precision as well as ultra-precision machines combined with an improvement in the dynamics. Most machines in high-technology areas require new drives for motion in several coordinates with positioning accuracy in the sub-micron area and fast dynamics. The paper presents a multi-coordinate drive developed at the Institute of Microsystems Technology, Mechatronics and Mechanics and illustrates the advantages of its new structure compared to conventional drive principles.

1 INTRODUCTION

In recent years many products appeared on the world market which have completely new properties and system parameters compared to traditional products. They are a result of the integration of electromechanics, microelectronics, hard- and software as well as optics.

Already in the early 70's the Japanese Ministry of Trade and Industry introduced the term "MECHATRONICS" for this particular field as a connection of *mechanics and electronics*. The multi-coordinate drive described below presents a typical mechatronic product.

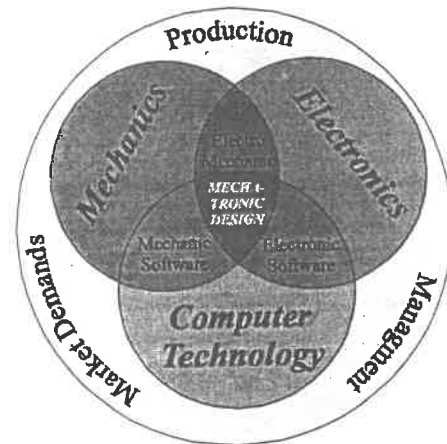


Fig. 1 Component integration of mechatronic Systems

2 STRUCTURES OF MULTI-COORDINATE DRIVE SYSTEMS

In most cases today's x-y-positioning systems generally consist of two serially coupled axes according to pick aback principle. The linear axes are usually fitted with a rotary motor (DC-servo motor, stepping motor) and a motion transmitter for generating the linear motion. In doing so (Fig. 2) the y-axis has to be carried by the x-axis. Beside the hiking loads created thereby the features of these multi-coordinate drives are determined by the clearance, the mechanical hysteresis and the elasticity of the essential motion transmitters.

These drives are limited in their positioning accuracy and speed not only because of the limited speed of motor spindle arrangements but also because of the high mass of coupling and guiding elements between the single axes.

The replacement of the motor-spindle arrangements by linear axes, however, gives only a small improvement. Substantial improvements can be obtained when the principle of functional integration in the area of mechanics and the principle of functional separation in the area of the automatic control is applied to the actuators of these so called multi-coordinate drives. The combination of

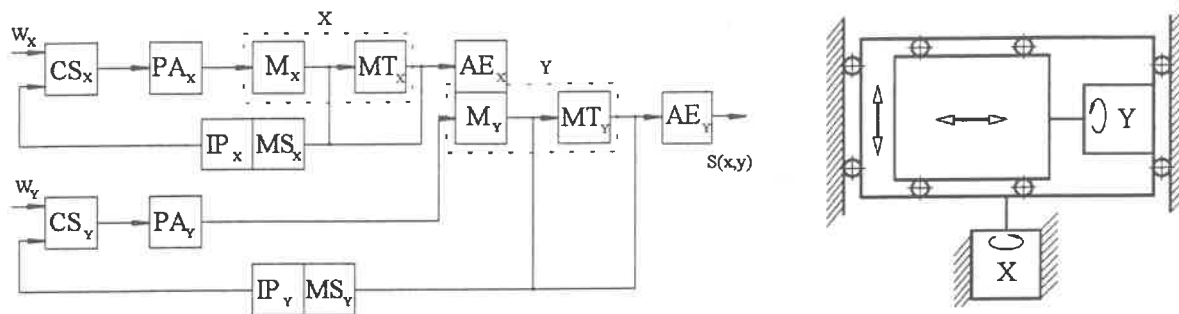


Fig. 2 Functional structure of multi-coordinate x-y-drives with serial arrangement of axes
 CS control system, PA power amplifier, M motor, MT motion transmitter, AE acting element
 MS measuring system, IP information processing

several single axis into *one integrated module* with a DOF of $F > 2$ eliminates the disadvantages of conventional systems mentioned above. Functional structure and symbolic description of such a drive is shown in Fig. 3.

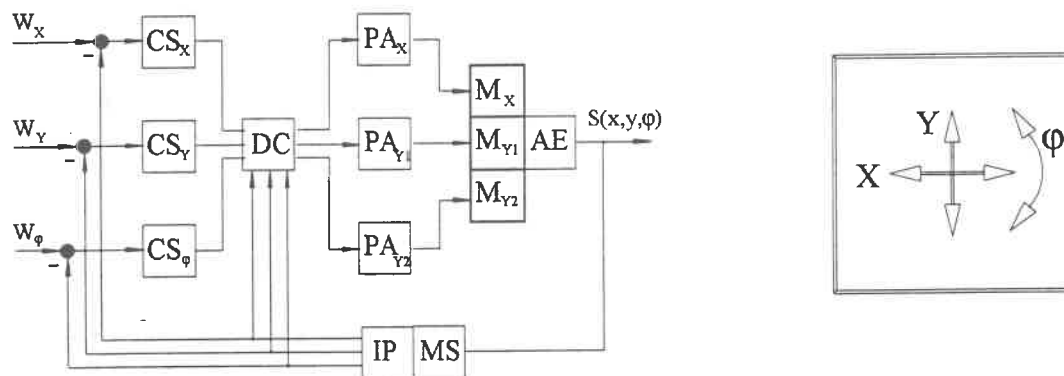


Fig. 3 Functional structure of multi-coordinate drives with integrated arrangement of axes
 CS control system, PA power amplifier, M motor, MT motion transmitter, AE acting element
 MS measuring system, IP information processing, DC decoupling controller

3 MEASURING SYSTEM

Integrated multi-coordinate drives can acquire extraordinary properties when if any mechanical rotary lock against the rotation of the actuator around the z-axis can be avoided. In this case the positioning accuracy in x-y-direction depends only on the quality of the measuring system applied. One possibility to eliminate the mechanical rotary lock is the guidance of the actuator in ϕ -direction also with electromagnetic forces. Then the drive has to be able to generate forces not only in the linear coordinates but also a torque around the vertical axis z (x-y- $\Delta\phi$ -motor). Furthermore a multi-coordinate measuring system has to be implemented which detects besides the linear

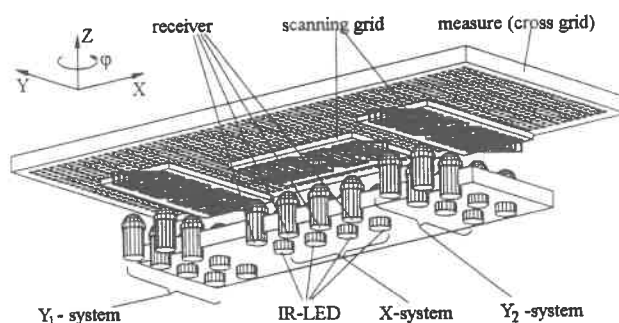


Fig. 4 x-y- $\Delta\phi$ -measuring system

coordinates (x, y) also the rotation of the actuator around the z-axis to a certain degree. The measuring systems for the linear coordinates (x, y), however, have to be insensitive to that rotation up to a certain degree. Fig. 4 shows a schematic representation of a suitable photoelectric x-y- $\Delta\phi$ -measuring system. It works according to the incident light technique. A cross grid serves as measure with a grid constant of $g=40\mu\text{m}$. The scanning plate carries groups of line grids, 4 for the x- and twice 4 for the y-coordinate. These line groups have the same grid constant and are shifted by $g/4$ each, similar to conventional linear incremental measuring systems.

In general such measuring systems especially those having small grid constants are highly sensitive to the relative twist of the cross grid and the scanning grid. However, by modifying the size of the scanning window and changing the dimensions of the optical system, this sensitivity can be reduced considerably, and the actuator can turn up to several degrees. Fig. 5 shows the decrease of the amplitude of the primary signals from the measuring system as a function of the relative twist angle of the scanning grid.

The decreasing amplitude of the signals from the measuring system has to be compensated by a special interpolation method. At the same time, a correction of the dc level, the phase shift and the error caused by the harmonics of the primary signals can be achieved with software to a certain degree.

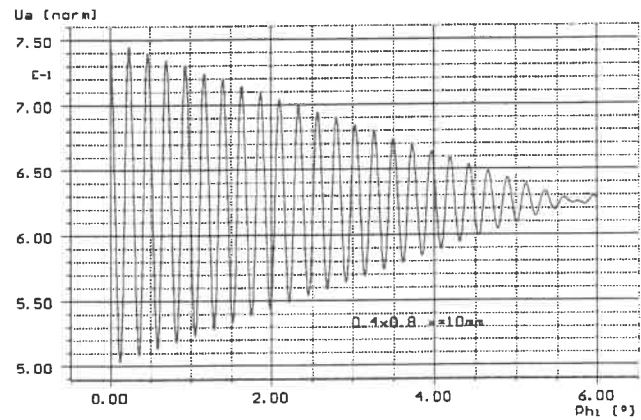


Fig. 5 Signal characteristic as function of angle ϕ

4 EXAMPLES FOR POSSIBLE MECHANICAL STRUCTURES

On principle there are two possibilities for the construction of a multi-coordinate electrodynamic drive - moving coils or moving magnets.

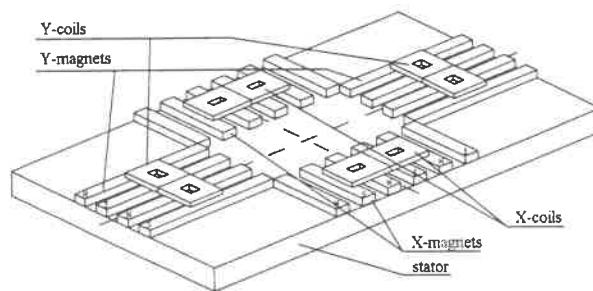


Fig. 6 Drive with moving coils

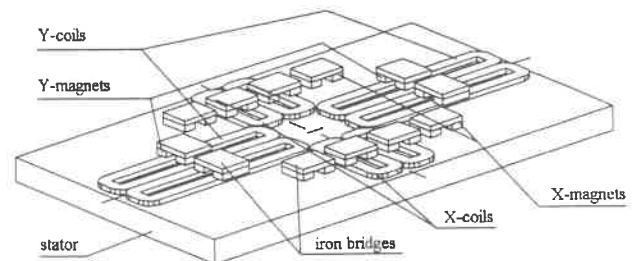


Fig. 7 Drive with moving magnets

Fig. 6 shows an integrated x-y- $\Delta\phi$ direct drive with moving coils. On a stator plate of iron flat permanent magnets are symmetrically arranged for both coordinates. The actuator carries a number of flat copper coils. The iron yokes which close the permanently excited magnetic circuits are not shown in the figure. Opposite coils generate the driving force (Lorentz force) along one coordinate. They are commutated according to the relative position of actuator and stator. In general

it will be advantageous for a highly dynamic drive to put the driving elements (coils or magnets) with the lower mass in the actuator. In case of moved coils this means short coils and long magnets. Fig. 7 shows the other way of generating an x-y- $\Delta\phi$ -motion with electrodynamic forces. On a stator plate pairs of long and flat copper coils are symmetrically arranged for each coordinate. 'Flying' magnets fixed in the actuator and closed to short circuits by iron yokes form the moving element. The guidance of the actuator in z-direction is favourably realised with air-bearings or magnetic levitation elements. Both demonstrated mechanical structures imply a number of advantages and disadvantages. Advantages of the second structure are

- a magnetic attraction between stator and actuator which permits an overhead operation of the drive and pre-loads the air bearings,
- a relatively small volume of the magnetic material needed,
- relatively light iron bridges and short magnetic flux ways,
- no disturbing forces on the actuator due to trailing cables for energy supply, and
- free accessibility of the actuator.

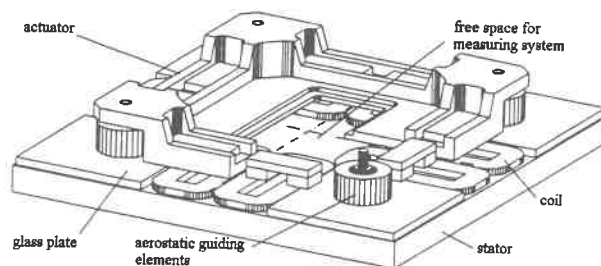


Fig. 8 CAD picture of realized drive

In order to create light actuators, a multiple array of driving elements for each coordinate is advisable. By doing so the moved iron volume

for closing the magnetic circuits can be reduced to a minimum. By having a current flow through the opposite coils of one coordinate in different direction a torque can be generated, and the position of the actuator can be controlled to be $\Delta\phi=0$.

Fig. 8 shows a newly designed drive system. Its structure corresponds to the system represented in Fig. 7. The actuator is guided on 4 air-bearings in the x-y-plane. The drive does not have any mechanical elements preventing the armature rotation around the z-axis.

The major advantage of this is, however, that the guidance system does not affect the accuracy of the three controlled axes anymore. The stage is guided electronically in those axes by the control system with the full accuracy of the measuring system. The mechanical guidance system takes only care of the so far uncontrolled z-axis and prevents rotation around the x- and y-axes.

The realised range of motion is $104 \times 104 \text{ mm}^2$. In the beginning of 1997 a drive with $208 \times 208 \text{ mm}^2$ will be available. In the centre of the actuator there is free space for mounting the measure of the measuring system. The scanning system (Fig. 4) is integrated in the stator.

5 Control of the drive

5.1 Control Hardware

The control of such a complex system requires a powerful hardware. Though control algorithms are inherently sequential the data acquisition and interpolation for the measuring system can be performed in parallel. Therefore, a transputer systems of seven processors and a PC as host is utilised (Fig. 9). Despite the parallel processing the limits of the system have been reached because of the comparatively small computing power of a single transputer processor. Therefore the system is currently being ported to a flexible DSP-hardware.

- P7: control output (1 ms)

5.3 Control Algorithm

A control concept called integral LQ-control with finite optimisation horizon is used [2]. It combines the advantages of state feedback control with the robustness and static disturbance cancellation properties of classical PID-control. Its structure is shown in Fig. 11.

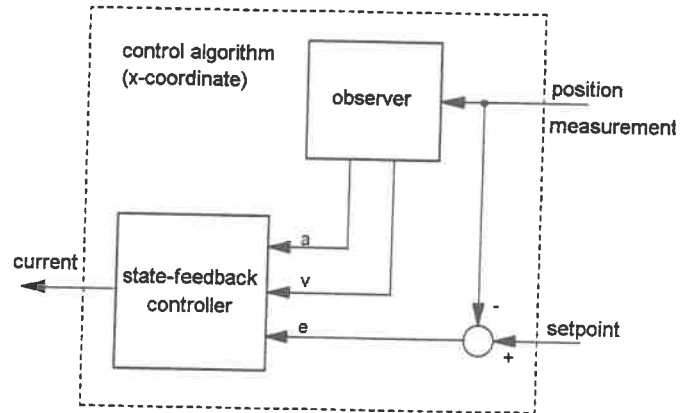


Fig. 11 Control structure

REFERENCES

- [1] Schäffel, Chr., Schober, U., Kallenbach, E. *Measuring system for x-y-Drives without Mechanical Elements for Preventing the Armatures Rotation Round the z-Axis*, Proc. PCIM '94 Nürnberg, pp 173.
- [2] Radtke, M. and Saffert, E. *Robust Adaptive LQG-Control under Consideration of the Identified Model Uncertainty (in German)*, Proc. 39th Int. Scient. Colloquium TU Ilmenau, Vol. 3, pp 723-730, 1994.
- [3] Schäffel, Chr., Saffert, E. and Kallenbach, E. *A new integrated multi- coordinate drive with anti-rotation lock through field forces (in German)*, Proc. of the 40th Int. Scient. Colloquium TU Ilmenau, Vol. 4, pp 469-475, 1995.
- [4] Saffert, E.; Schäffel, Chr.; Kallenbach, E. *Control of an Integrated Multi-coordinate Drive*, Proc. Mechatronics '96 conference, Portugal, 1996.

DESIGN OF A LONG RANGE FAST TOOL SERVO SYSTEM USING MAGNETIC SERVO LEVITATION

Brent A. Stancil
Harris Corporation
Government Communication Systems
Melbourne, FL 32902

Paul I. Ro
Precision Engineering Center
North Carolina State University
Raleigh, NC 27695-7918

1 INTRODUCTION

Magnetic Servo Levitation (MSL) [1] has been proposed as the basis of a Long Range Fast Tool Servo System (Figure 1). A system has been designed that is capable of a 1 mm range of travel, 300 Hz operating bandwidth, and nanometer resolution [2] (Figure 2). MSL utilizes the attractive forces of electromagnets to provide the actuating force for the system. Therefore, the system is open loop unstable and requires closed loop control to make any testing viable. Also, due to the large mass, long range, and decaying behavior of electromagnetic forces, the system was designed to have a natural frequency of 300 Hz, or the same as the desired bandwidth. Therefore, closed loop control will be necessary to provide the 300 Hz bandwidth. This abstract will attempt to show that the design is adequate and can be controlled.

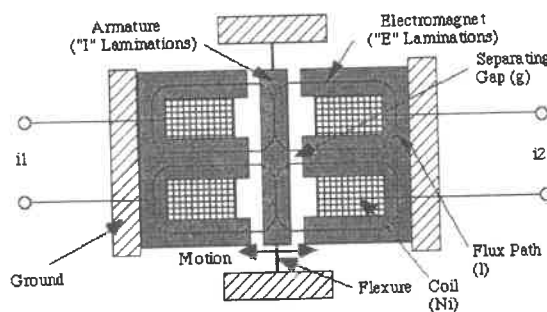


Figure 1: MSL Schematic

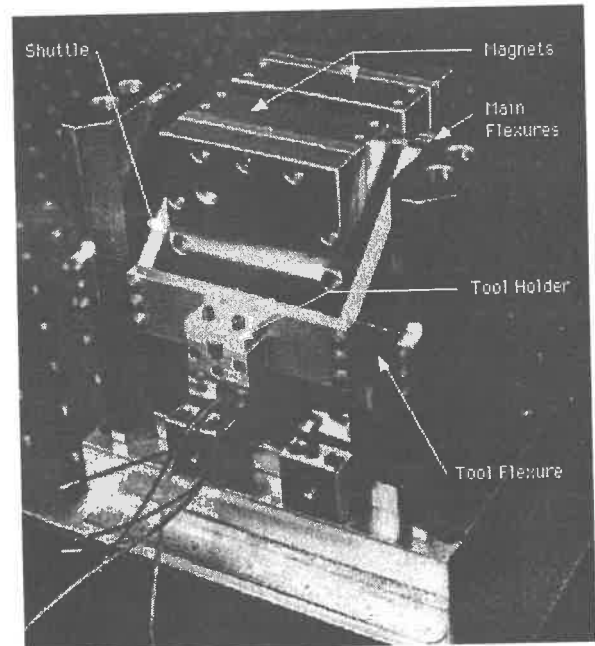


Figure 2: Long Range Fast Tool Servo

2 CONTROL

In order to provide closed loop control of nanometer resolution over a 1 mm range, a laser system had to be implemented into the final design. A Zygo ZMI-1000 system was chosen, because of its 2.47 nm resolution, capability of handling speeds up to 1 m/s, and its competitive cost of about \$20,000. A version of a plane mirror interferometer was chosen that uses a focusing optic rather than a double pass configuration. This is required to accommodate the velocity requirement. A schematic of the system is shown in Figure 3.

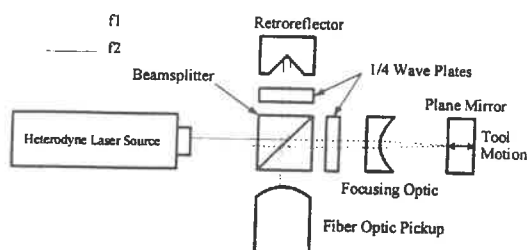


Figure 3: Laser Schematic

Two different control algorithms were used to show the effect of control on the open loop system. First, a robust time delay controller [3] was implemented. Its operating principle was that the input would be shaped in a way to cancel the dynamics of the system, or it add zeros to cancel the poles of the system. The controller, $D(z)$, has the form, in the digital domain,

$$D(z) = \frac{A_1 z^{k+1} + 1}{(A_1 + 1)z^{k+1}} \quad (1)$$

where A_1 is found by

$$A_1 = \exp\left(\frac{\zeta\pi}{\sqrt{1-\zeta^2}}\right) \quad (2)$$

and k is found by

$$k = \frac{T}{T_s} \quad (3)$$

where T is the time delay found by

$$T = \frac{\pi}{\omega\sqrt{1-\zeta^2}} \quad (4)$$

where ω is the natural frequency of the system, ζ is the damping ratio of the system, and T_s is the sampling rate of the controller.

The second control algorithm used was pole placement [4]. This algorithm operates on the principle of actually moving the system poles, and was investigated for the reason of extending the bandwidth.

3 TESTING

Testing was done on four different systems. First, testing was done open loop, only dividing the reference signal by the open loop gain. Secondly, the robust time delay controller was implemented. Thirdly, the pole placement routine was tested. Finally, the robust time delay was used to filter the reference signal to the pole placement controller. Bode plots of each of these are shown in Figures 4-7. Two time response tests were done for each configuration as well. Figures 8-11 depict the results for a 125 μm 20 Hz sine wave command and Figures 12-15 depict the response for a 200 nm 20 Hz square wave command. The 125 μm sine wave shows the ability of the system to achieve longer ranges than previous fast tool servo systems. This is also the limit of stability for the open loop case so that a comparison could be made with the controlled systems. The 200 nm sine wave shows the ability for the system to handle nanometer scale commands. These are typical amplitudes for error correction. The 20 Hz frequency corresponds to a once per revolution variation of the part being manufactured for a spindle speed of 1200 rpm.

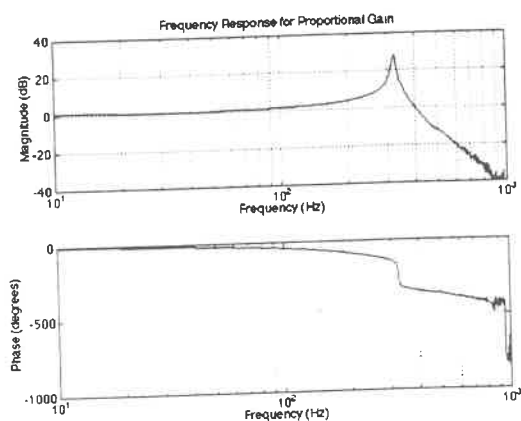


Figure 4: Open Loop Bode Plot

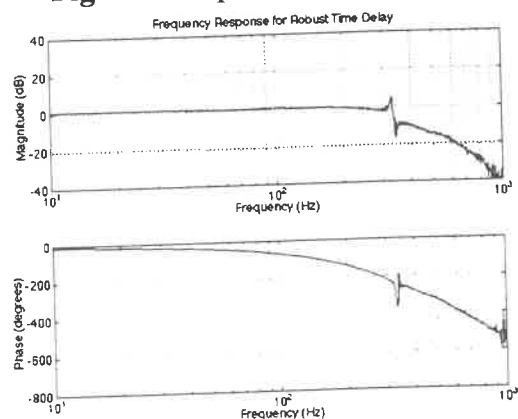


Figure 5: Robust Time Delay Bode Plot

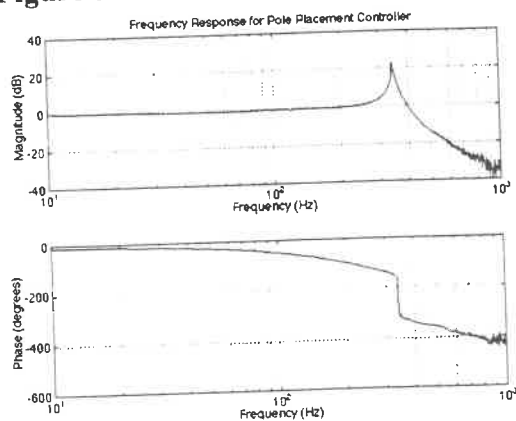


Figure 6: Pole Placement Bode Plot

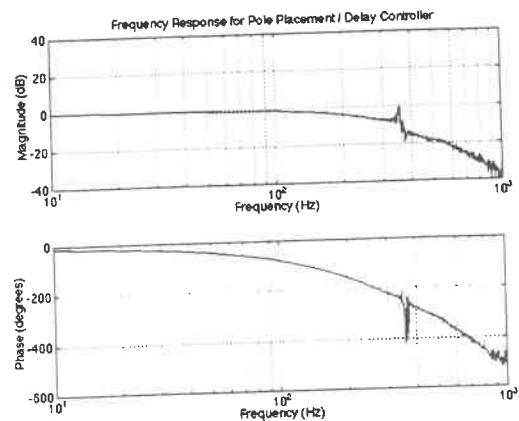
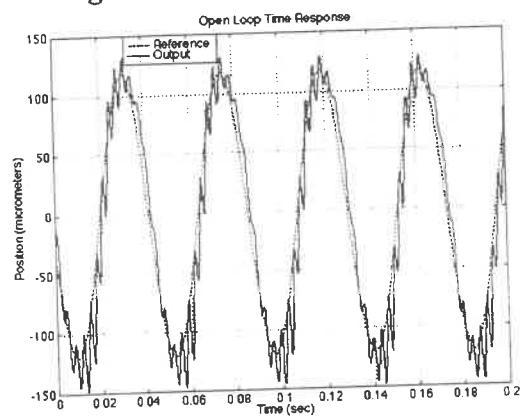
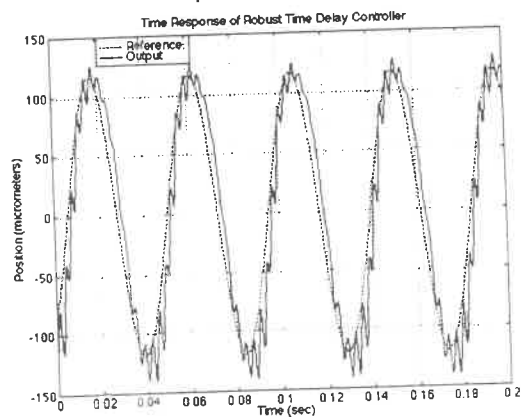


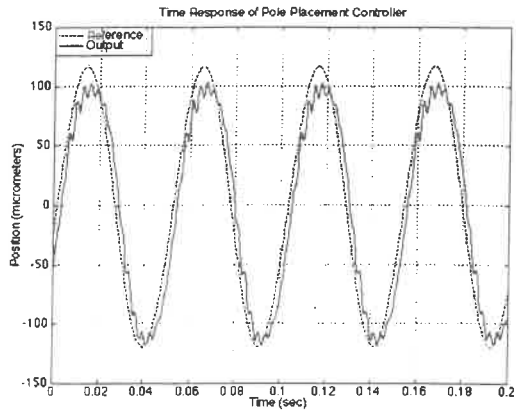
Figure 7: Combined Bode Plot



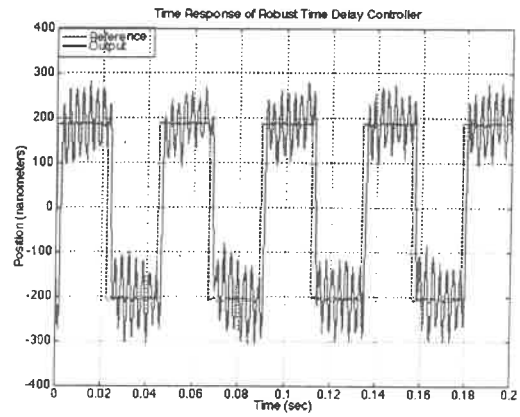
**Figure 8: Open Loop
125 μ m Sine Wave**



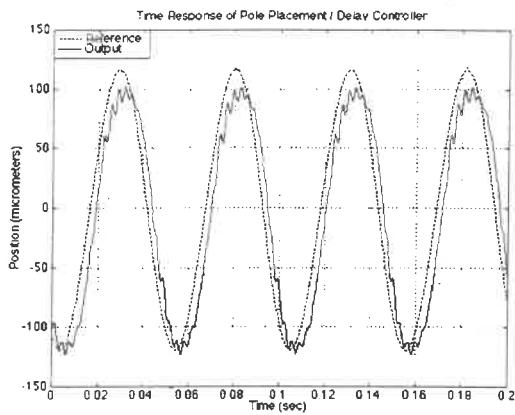
**Figure 9: Robust Time Delay
125 μ m Sine Wave**



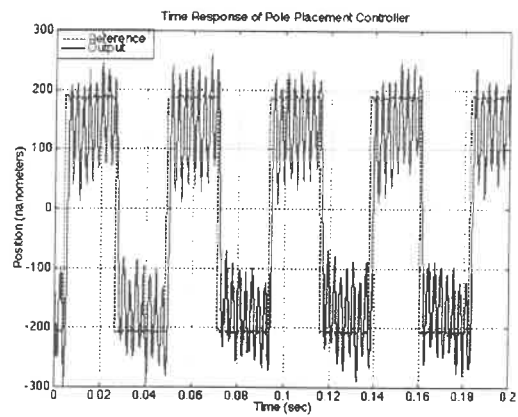
**Figure 10: Pole Placement
125 μm Sine Wave**



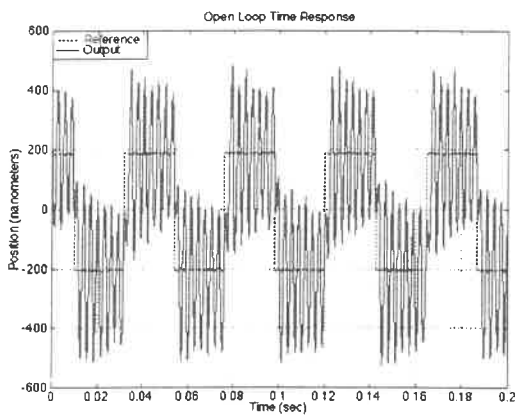
**Figure 13: Robust Time Delay
200 nm Square Wave**



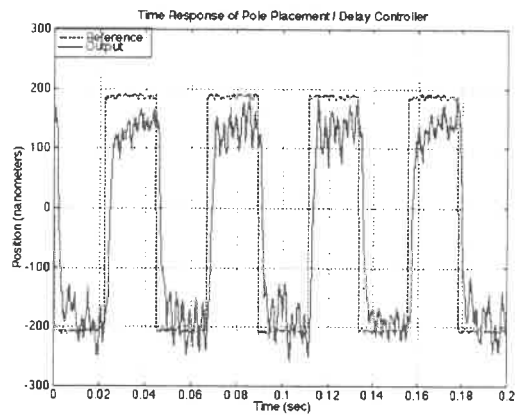
**Figure 11: Pole Placement / Time Delay
125 μm Sine Wave**



**Figure 14: Pole Placement
200 nm Square Wave**



**Figure 12: Open Loop
200 nm Square Wave**



**Figure 15: Pole Placement / Time Delay
200 nm Square Wave**

The open loop bode plot (Figure 4) shows a mechanical resonance at about 325 Hz with about a 28 dB gain. The robust time delay (Figure 5) decreased this gain to about 5 dB. However, some phase was added to the system, which is expected since this is a time delay controller. The pole placement routine extended the bandwidth of the system by flattening the 0 dB region for a higher range of frequencies (Figure 6). However, some concern is raised since the resonant peak was not damped dramatically with the real poles that were designed for. This can be attributed to the nonlinearity of the system. The phase did not differ dramatically from the open loop system. The combination of the two controllers (Figure 7) provided a good frequency response. The resonant peak had been damped to 0 dB. However, the bandwidth was not quite as good as the time delay or pole placement. The phase was better than the case of the time delay by itself.

The open loop time response of the 125 μm sine wave (Figure 8) shows that the system is capable of tracking a command signal. However, there is an undesirable 325 Hz oscillation on top of the 20 Hz command signal. The magnitude of this oscillation is reduced somewhat by the robust time delay (Figure 9). Pole placement improves on this ringing to an even greater extent (Figure 10). The combination of the two controllers produces the cleanest signal (Figure 11).

Finally, the open loop time response of the 200 nm square wave was extremely noisy, with about 200 nm of unwanted oscillation (Figure 12). By adding the robust time

delay, this oscillation was reduced to about 80 nm (Figure 13). However, the pole placement controller (Figure 14) was not as effective as it was with the 125 μm sine wave, and only reduced the oscillation to about 80 nm as well. However, the combination of the two controllers reduced this oscillation to about 40 nm (Figure 15).

CONCLUSIONS

By the addition of low order control algorithms to the Magnetic Servo Levitation based Fast Tool Servo System, bandwidth could be increased and the resonant peak could be damped considerably. However, there is some remaining effects of the resonant peak due to the nonlinearity of the system exciting this frequency. Higher order controllers should be able to help this system achieve the desired design goals.

REFERENCES

- [1] Tsuda, M., and T. Higuchi. Design and control of magnetic servo levitation. *Report of the Institute of Industrial Science*. The University of Tokyo, March 1992, Vol. 37, No. 2, pp. 137-205.
- [2] Stancil, B. A., H. M. Gutierrez, and P.I. Ro. Design of a Long Range Fast Tool Servo System Using Magnetic Servo Levitation. *Proceedings from ASPE 1995 Annual Meeting*, pp. 301-304.
- [3] Singh, T. and S. R. Vadali. Robust Time-Delay Control. *Journal of Dynamic Systems, Measurement, and Control*. June 1993, Vol. 115, pp. 303-306.
- [4] Phillips, Charles L. and H. Troy Nagle. *Digital Control System Analysis and Design*, 3rd Edition. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1995.

A neural network-based real-time compensation technique for frequency output sensor arrays

G. Y. Tian

Z. X. Zhao

R.W.Baines

(School of Engineering, University of Derby , Kedleston Road, DE22 1GB, UK)

1. Introduction

Global and local networks have become the major computing platform for manufacturing systems. It becomes important to get the information from measurement sensors and to log them into computer based systems and/or networks. To do this sensors need to be compact, high-accuracy, safe, reliable and "smart and/or intelligent" enough to be integrated with different computing environments. To achieve this, smart and intelligent sensors have been researched and developed. Following this way, there is still a long way to go. Many techniques should be considered. They are as following:

- A. New sensor theories, new materials and techniques based on high technology were adopted for developing sensors and their measurement and control system.
- B. It is now widely accepted that new forms of sensors have improved performance over existing devices and are compatible with the extensive number of computer controlled system. With the development of the computer application, sensor compatibility with computer based system become emergency.

Due to the second opinion discussed above, a new approach to frequency output sensor arrays based on artificial neural network is proposed. The self-calibration/compensation technique including nonlinearity modelling recognition and coefficient evaluation/compensation, drift correcting readout during changing environmental conditions is discussed in the paper. There were three kinds of courses for self-calibration/compensation involved in the paper. One is for the non-linearity. The others are for stability relating to the change of temperature and humidity for both short and long term drift. The research work was taken into two steps.

Firstly, we designed and built a system linking frequency output sensors into PC. Based on the system, the data acquisition was provided. Different algorithms of neural networks were used for nonlinearity modelling recognition , coefficient estimation and compensation. Different sensors supported by different neural networks were discussed. The flexibility and accuracy of the self-calibration/compensation were validated.

Secondly, the self-compensation/compensation is imbedded in microprocessor-based sensor arrays. We used EEPROM instead of conventional ROM [1-2]. By using neural networks trained by a simple neural algorithm, the unknown coefficients of sensors can be estimated accurately and stored in EEPROM for real-time signal acquisition. When environmental conditions such as temperature and humidity are changed or sensors are to be renewed, the coefficients can be updated in real-time. From the coefficients of

compensation for long term drift, the quality information can be collected and the life time of sensors can be evaluated.

Finally as a conclusion, sensors with frequency output is an effective approach for resolving the "bottleneck" in computer application in measurement and control due to their the direct digital method and interface technique [3]. The approach presented in the paper will increase the intelligence of the sensors.

2. Frequency output sensor

Sensors transfer physical values into frequency especial digital pulse signal directly. This kind of sensors is called sensors with frequency output or frequency output sensors. The diagram of frequency output sensors was shown in Fig. 1. In this sense, many methods have been reported for the frequency converters [3]. We have designed electric oscillators for mechatronic sensors with frequency output. The stability of oscillators are the important factor for these sensors. The quartz oscillator is known for its dependence on temperature. Therefore, design quartz oscillators for different input quantities of sensors were used in our research.

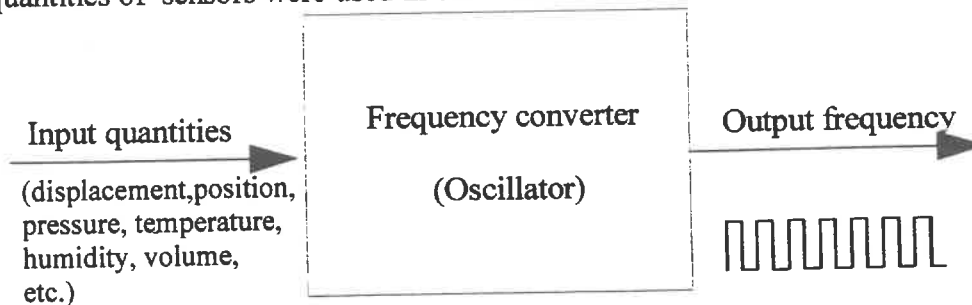


Fig. 1 Block diagram of the frequency output sensors

The advantages of research and development this kind of sensors are shown as following:

- (1) sensors with frequency output transfer physical values into digital values directly. In this way, it is very easy to realise the interface from sensors and computer. Further, it can be reduced the design work for designer to design and adjust analogous circuit such as AC/DC amplifiers, filters. Therefore, using this kind of sensors do not require expensive A/D converter for computer interfaces.
- (2) Since measuring frequency is a cumulative process such as output pulses are summed or integrated over a fixed period of time by numerical method using a microprocess or microcomputer or using IC unit, unwanted signal are inhibited. And the frequency output signals make the sensors insensitive to temperature and humidity change. So measuring frequency can achieve higher accurate. On the other hand, digital frequency signal have better anti-interference ability. It is operable in remote, hostile environment. Therefore, it can use for network or complicated measurement and control system especially under mal-condition environment.
- (3) The sensors with frequency out can transfer analogue information into domain and can be digitised by a micro control directly. It can be helpful digitisation, integration and intelligence for sensors such as smart sensor and mechatronics devices. It can integrate

sensors and computer IC together, behaving functions about processing information, self-calibration, self-linearation, self-compensation, self-diagnositious and so on.

For example, an approach of intelligent system for frequency output sensors was shown in Fig.2. As we knew, sensor performances are generally decided by sensor drift including temperature drift and long term drift, non-linearity and resonant time. Resonant time will be discussed later. Relating to the stability and linearity of sensors, design high quality of the frequency converters is one resolved method, on the other hand, an intelligent system shown in Fig. 2 is transmitting the frequencies by the sensors to a microcomputer or embedding to a microprocessor system, via multi-cycle synchronous counter .

Electrically Erasable Programmable Read only Memory (EEPROM) is used for saving programme code, data for compensation and non-linearity calibration and pattern recognition. Based on the intelligent system, numerical computation such as fuzzy logic, neural network and genetic algorithms could be integrated. The self-calibration, self-linearation and self-compensation for draft in the intelligent system are being studied in our research. By incorporating a single or multiple low-cost, uncalibrated frequency output sensor(s) into a software feedback loop, a self-normalised calibration/compensation table for each sensor can be generated and stored in EPROM for later use in real-time signal acquisition.

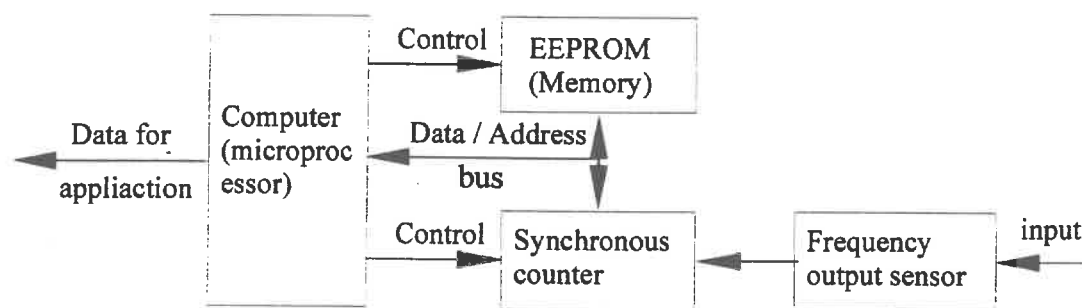


Fig. 2 Schematic diagram of an intelligent system for frequency output sensors

- (4) This direct digit method is ideal for digital update of a lot of traditional dial gauges such as flow measurement and electricity measurement and further to design as a computer controlled network.
- (5) Research and development this kind of sensors can simplify the sensor technology. Research only needs to pay attention how to design frequency output for sensors and interface for measuring frequencies.

3. Experimental test

An experimental test rig of an eddy current noncontact sensor with frequency output has developed at Derby university. The measured variable is displacement and the sensor output is digital frequency. The original relationship between the input and output quantity are shown as figure 3. In the experimental test, N stands for the test numbers for the same target at different positions with each other 10 mm. The output frequency can be got by a digital frequency meter. It is very clear that the nonlinearity is serious.

Based on the excellent long-term and temperature stability of the rig tested in the experiment, different compensation and calibration were taken. Using neural network software, Adaline and Madaline network and adaptive resonance theory networks have been used in the nonlinear compensation. The performances by the intelligent compensation are significant. Table 1 was shown the sensor performances compared with the traditional sensor.

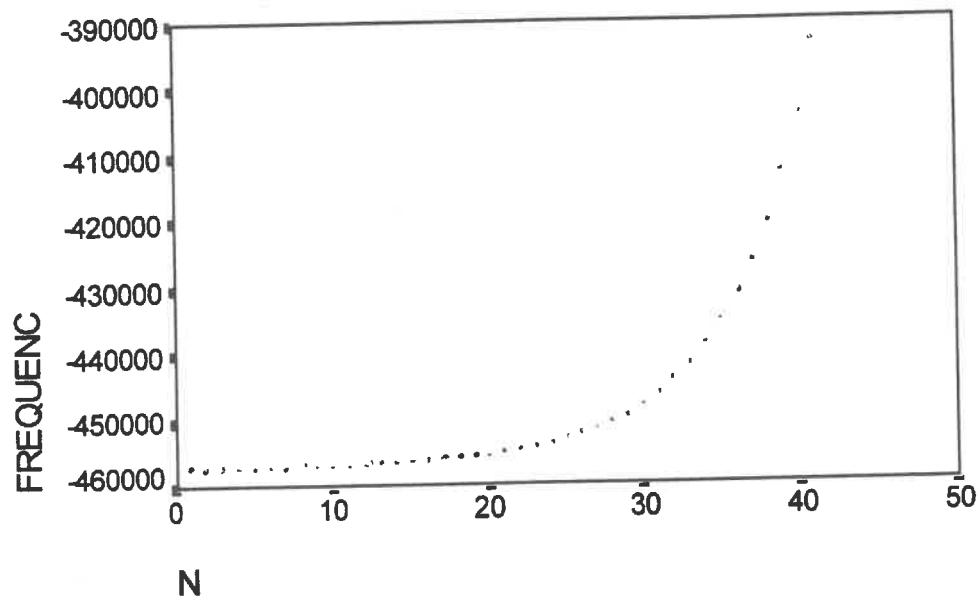


Figure 3 The frequency output for an eddy current sensor

Specification	New eddy current sensor	Conventional eddy current sensor
Range mm	0-2	0-2
Linearity %FS	0.1	0.25
Resolution μm	0.1	1
Temperature stability $\mu\text{m}/^{\circ}\text{C}$	0.2	2

Table 1 - Comparison table for technical specifications of non-contact displacement eddy current sensors

4. Conclusion

With ever shrinking tolerances and reduction in manufacturing variability, pressure is increasing for the measuring and gauging industry to continually improve their own capabilities. The frequency output sensor approach can improve its stability and accuracy. The neural network-based real-time compensation technique has a potential future in the field of instrumentation and measurement. Not only can it improve system performance, but it is easily to be integrated into the existing quality assurance system such as manufacturing , monitoring and maintenance systems

5. References

- [1] Jagdish Chandra Patra, Ganapati Panda, and Rameswar Baliarsingh, " Artificial Neural Network-Based Nonlinearity Estimation of Pressure Sensors ", IEEE transactions on instrumentation and measurement, Vol. 43, No. 6. Dec. 1994.
- [2] Paul T.Kolen, " Self-calibration/Compensation Technique for Microcontroller-Based Sensor Arrays ", IEEE transactions on instrumentation and measurement, Vol. 43, No. 4. Aug. 1994.
- [3] Tian Guiyun, Yan Bin, " A New Fringe-field Capacitive Sensor ", The proceedings of 10th annual meeting for American Society for precision Engineering (1995 ASPE).

Development of NC Program Evaluator for Higher Machining Productivity

Xiren Yan, Keiichi Shirase, Masatoshi Hirao and Takeshi Yasui
Faculty of Technology, Kanazawa University
Kanazawa, Japan 920

1. Introduction

Numerical control machines, especially machining centers (MC), have been widely used in the precision and high speed cutting of complicated workpieces such as molds and dies. Their makers, as well as users, are focusing their attention on the machining productivity. Thus far, the productivity has been evaluated by the operational time and preparatory time generally. However, the productivity is greatly affected by NC programs used. Substantially, it should be evaluated by analyzing the NC programs, specifically the cutting conditions, tool paths, and cutting times in them. In other words, the NC programs should be analyzed internally. Thereby the quality of the NC programs will be evaluated clearly, and by employing the self-constructed expert system on cutting conditions based on the cutting know-how, the NC programs will be optimized.

Up to now, some controllers of machining centers and commercial software are available for part of the job. However, those only give the information on the cutting time, and far from evaluating and analyzing the productivity. Besides, due to the dependence of NC codes on controllers, and some optional codes added to NC code systems, no software for this purpose will apply for every kinds of NC code system. Moreover, as the cutting speeds and precision increase, a large number of tiny blocks, or the blocks where the moving distance is much shorter than that of the conventional case, appear in the NC programs, so that the actual feedrate is much lower than the commanded one. Consequently, the ideal cutting speeds and precision will not be obtained, and that will result in the calculations of cutting times based on command feedrates and moving distances far from actual ones. This causes the calculations much more difficult than imagined. In consideration of the above factors, an NC program evaluator has been developed for analyzing NC programs or evaluating cutting conditions, tool paths and operational times. With the evaluator, even before the actual cutting, the tool paths can be displayed and the times of feed motion and rapid motion, etc. can be calculated. Furthermore, the quality or productivity can be evaluated and expected to be improved. The overall block diagram evaluating and optimizing NC programs is shown in Fig. 1¹.

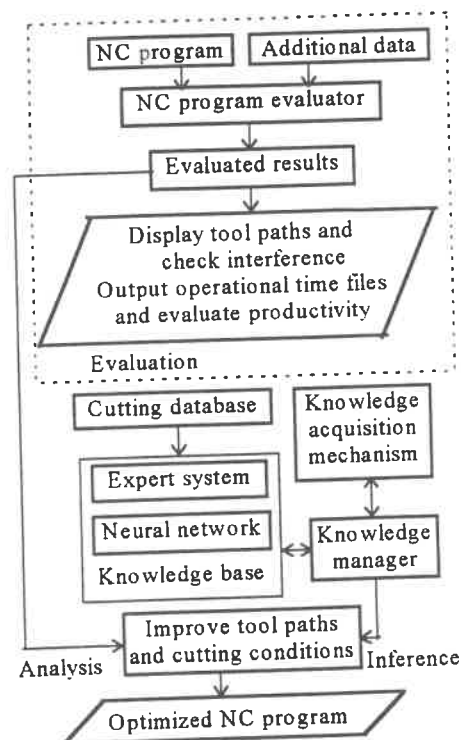


Fig.1 Overall block diagram of the system

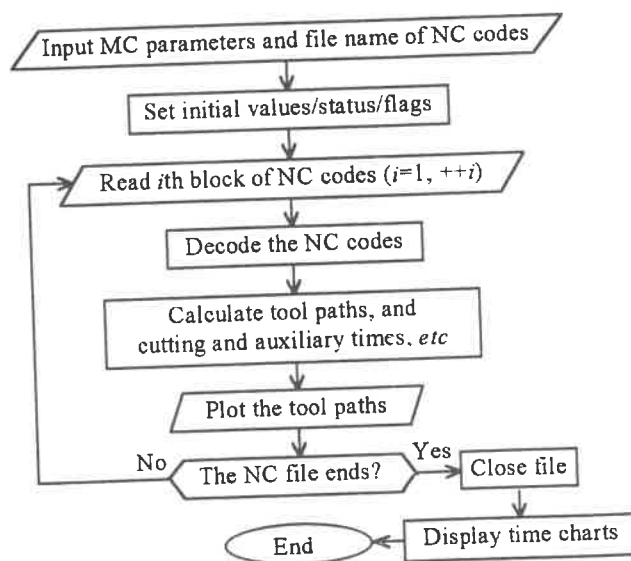


Fig.2 General flow chart of the evaluator

2. NC Program Evaluator

The NC program evaluator works by decoding NC codes such as G and M in an NC program block by block, as shown in Fig.2. As decoding goes, the tool

paths are displayed and the feed times are calculated.

2.1 Tool paths

By decoding NC codes, the relevant coordinates of the tool are calculated and tool paths are drawn corresponding to G00, G01, G02, G03, etc. Furthermore, the quality of the tool paths can be evaluated.

2.1.1 Rapid feed (G00, G28, G29 and G30)

In the rapid feed, the tool path in one block is generally not a straight line but a broken line or sectional line, because the x, y, z axes in an MC are controlled separately. Therefore, in addition to the commanded dimensions (X, Y, Z etc.), the coordinates of the midpoints where the tool traverses should be calculated. These can be obtained by comparison of the moving times in x, y, z axes. Figure 3 shows the calculation procedure of the midpoints M_1, M_2 .

Letting S, D denote the start-point and destination point in the block respectively, then the positioning tool path is $S - M_1 - M_2 - D$.

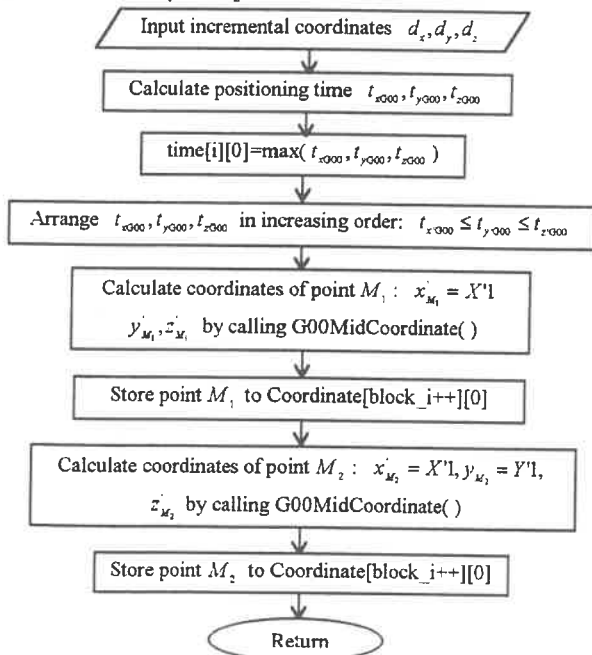


Fig.3 Calculations of the midpoints in the rapid feed

2.1.2 Circular cutting (G02 and G03)

In the circular cutting, the center of the arc or vertices of its boundary rectangle should be calculated beforehand. Given the start-point $S(x_s, y_s)$ and the destination point $D(x_d, y_d)$, and the radius R , from the geometrical calculation, the coordinates of the center of the arc are obtained:

$$x_c = (x_s^2 - x_d^2 + y_s^2 - y_d^2 + 2(y_d - y_s)y_c) / 2(x_s - x_d)$$

$$y_c = (y_s^2 - y_d^2 + x_s^2 - x_d^2 + 2(x_d - x_s)x_c) / 2(y_s - y_d)$$

$$x_c = (x_s^2 - x_d^2 + y_s^2 - y_d^2 + 2(y_d - y_s)y_c) / 2(x_s - x_d)$$

$$y_c = (y_s^2 - y_d^2 + x_s^2 - x_d^2 + 2(x_d - x_s)x_c) / 2(y_s - y_d)$$

where + is for G02 and - is for G03.

2.2 Cutting times

The operational time for machining a workpiece includes the cutting (feed) time, positioning (rapid feed) time, and ac/deceleration time of the spindle, etc.

In the cutting feed motion, the cutting time will be much different from the actual one in some cases if it is calculated just based on the command feedrates (F code) and coordinates (X, Y, Z etc.) because the feedrates and accelerations are limited by some optional codes and accelerations. Besides, when the moving distance for each block is too short, the actual feedrate cannot reach to the command feedrate, especially in the high speed cutting. In the following sections, the calculations of the cutting time are derived for the feedrates of both the exponential and the linear functions.

2.2.1 Feedrate of exponential function

The instantaneous feedrate in an axis can be derived from the block diagram of the CNC controller. For example, the feedrate in x axis is expressed as

$$v_x(t) = \frac{v_{x0} - v_{x1}}{T_1 - T_2} (T_1 \exp(-\frac{t}{T_1}) - T_2 \exp(-\frac{t}{T_2})) + v_{x1} \quad (2)$$

Integrating Eq. (2) yields the moving distance:

$$x(t) = \frac{v_{x1} - v_{x0}}{T_1 - T_2} (T_1^2 \exp(-\frac{t}{T_1}) - T_2^2 \exp(-\frac{t}{T_2})) + v_{x0}(T_1 + T_2) + v_{x1}(t - T_1 - T_2) \quad (3)$$

where v_{x0}, v_{x1} are the components of the initial and commanded feedrates in x respectively; T_1, T_2 are the time constants of the ac/deceleration and positioning system respectively. T_2 is also called the reciprocal of the position loop gain. For FX-5 (Matsuura Machinery Corp.), $T_1 = 32\text{ms}$ and $T_2 = 33\text{ms}$.

With the approximation $\exp(x) \approx 1 + x$ ($x \ll 1$), and in consideration of the final feedrate of the current block or the initial feedrate of the next one, from Eq. (3) the cutting time can be estimated as

$$t = (d + (T_1 + T_2)(2f_0 - f_{0S} - f_{0D})) / f_0 \quad (t \gg T_1) \quad (4)$$

where d is the moving distance in the current block; f_0 is the command feedrate or its limit; f_{0S}, f_{0D} are the initial and final feedrates in the block respectively.

When $f_{0S} = f_{0D} = 0$, Eq.(4) becomes

$$t = d / f_0 + 2(T_1 + T_2) \quad (t \gg T_1) \quad (5)$$

Equation(5) implies that the time calculated just by the commands F, and X, Y, Z etc. should be compensated by adding $2(T_1 + T_2)$ in this case.

Equation(4) can be used to calculate the cutting time

in the case of the exponential feedrate because t has a negligible error when $d > 4T_1 f_0$ and $T_1 \approx T_2$.

When $d \leq 4T_1 f_0$, the approximation of t derived from Eq.(3) will be a transcendental function. Even though it can be fitted with $d = At^2$ based on some numerical analyses, here for simplicity, Eq.(4) is still used for this case and the absolute error can be neglected.

2.2.2 Feedrate of linear function

In the case of linear feedrates, the relations of feedrate and time can be divided into three cases from the moving distance, feedrates and accelerations. Suppose the feedrate at time t in the current block is denoted by $f(t)$ and the command feedrate in the next block by f_1 . We will derive the formulas of the feed time corresponding to each of the cases. Note that the following formulas are relative to x axis, but the other axes are similar.

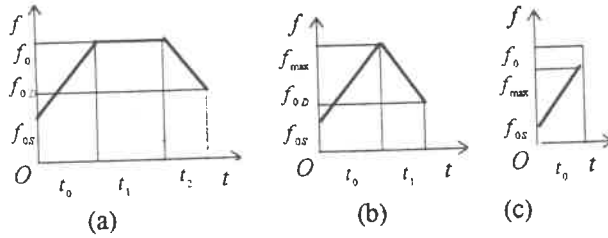


Fig.4 Feedrates in linear function

(1) Time estimation for the current block

Case 1: There exists time t where $f(t) = f_0$ (Fig.4(a))

When the following condition is satisfied

$$d_x' = |f_{0x}^2 - f_{0sx}^2 / 2a_{0x} + |f_{0x}^2 - f_{1x}^2 / 2a_{1x} < d_x \quad (6)$$

the maximum feedrate f_{maxx} in x axis reaches to the x component of the command feedrate f_{0x} (refer to Fig.4(a)). In this case, $f_{0Dx} = f_{1x}$, and the feed time in x axis is calculated by

$$t_x = |f_{0x} - f_{0sx} / a_{0x} + |f_{0x} - f_{1x} / a_{1x} + (d_x - d_x') / f_{0x} \quad (7)$$

where f_{0x}, f_{1x} are the x components of the command feedrates of the current and next blocks respectively; a_{0x}, a_{1x} are the x components of the initial and final feed ac/deceleration of the current block respectively; f_{0sx}, f_{0Dx} are the x components of the initial and final feedrates of the current block respectively; d_x is the commanded moving distance in x axis.

Case 2: There does not exist t where $f(t) = f_0$ but $f_{0D} = f_1$ exists (Fig.4(b))

When neither Eq.(6) nor the following equation is satisfied

$$f_{maxx} = \begin{cases} (f_{0sx}^2 + 2 \min(a_{0x}, a_{1x}) d_x)^{1/2} < f_{1x} (f_{0sx} \leq f_{1x}) \\ (f_{0sx}^2 - 2 \min(a_{0x}, a_{1x}) d_x)^{1/2} > f_{1x} (f_{0sx} > f_{1x}) \end{cases} \quad (8)$$

f_{maxx} does not reach to f_{0x} , but f_{0Dx} reaches to f_{1x} of the next block, i.e., $f_{0Dx} = f_{1x}$ (Fig.4(b)). In this case,

$$f_{maxx} = ((f_{0sx}^2 + f_{1x}^2 a_{0x} / a_{1x} \pm 2 a_{0x} d_x) / (1 + a_{0x} / a_{1x}))^{1/2} \quad (9)$$

where sign $-$ is used when $f_{0sx} > f_{0x}$ and the feed time

$$t_x = |f_{maxx} - f_{0sx} / a_{0x} + |f_{maxx} - f_{1x} / a_{1x} \quad (10)$$

Case 3: There does not exist t where $f(t) = f_0$ and $f_{0D} = f_1$ does not exist either (Fig.4(c))

When Eq.(6) is not satisfied but Eq.(8) is, we have $f_{0Dx} = f_{maxx}$, and the feed time

$$t_x = |f_{maxx} - f_{0sx} / \min(a_{0x}, a_{1x}) \quad (11)$$

where f_{maxx} is calculated by Eq.(8).

The feed time $t = \max(t_x, t_y, t_z)$.

(2) Time compensation for the last blocks

In the case of Fig.4(c), when the current block is the last of the consecutive moving blocks, i.e., $f_{1x} = 0$ and $f_{0sx} > \sqrt{2a_{0x}d_x}$, letting the final feedrate of the current block $f_{0Dx} = f_{1x}$, then the initial feedrate f_{0sx} is changed to $f_{0sx}' = \sqrt{2a_{0x}d_x}$, and thus the variation of the initial feedrate $\Delta f_{0sx} = f_{0sx} - f_{0sx}'$ is produced. Therefore, the feed time of the block should be compensated in corresponding to its feed cases. Differentiating Eqs. (7), (10) and (11) yields

$$dt_x = |(1 + f_{0Dx} / f_{0x}) df_{0Dx} / a_{1x}| \text{ (Fig.4(a))} \quad (12)$$

$$dt_x = |df_{0Dx} / a_{1x}| \text{ (Fig.4(b))} \quad (13)$$

$$dt_x = |(1 - f_{0Dx} / f_{0sx}) df_{0Dx} / a_{0x}| \text{ (Fig.4(c))} \quad (14)$$

Substituting df_{0Dx} in Eqs. (12)-(14) with Δf_{0sx} yields the approximate value $\Delta t_x \approx dt_x$.

For the consecutive tiny blocks, the compensation should be for all the blocks. From Eq.(14) and $df_{0sx} = f_{0Dx} df_{0Dx} / f_{0sx}$, it is recurred and summed that

$$\Delta t_x = \sum dt_x = (1 - f_{0Dx} / f_{0firstx}) df_{0Dx} / a_{0x} = f_{0Dx} / a_{0x} \quad (15)$$

where $f_{0firstx}$ is the initial feedrate of the first block.

(3) Ac/deceleration of circular cutting

In the circular cutting, the resultant ac/deceleration is used instead of that of a particular axis. The resultant ac/deceleration $a_t = df / dt$ and satisfies for the cutting in xy plane (G17)

$$\begin{cases} |a_x| = |\pm a_f \sin \alpha - f^2 \cos \alpha / r| \leq a_{\max x} \\ |a_y| = |\pm a_f \cos \alpha - f^2 \sin \alpha / r| \leq a_{\max y} \end{cases} \quad (16)$$

where the upper sign is for G03 and the lower for G02. α is the angle between the radial direction and x axis; f is the instantaneous resultant feedrate; r is the radius of the arc; $a_{\max x}$, $a_{\max y}$ are the maximum feed accelerations in x and y respectively.

In most cases, an NC program used for cutting molds or dies is composed of a great number of consecutive tiny blocks, and thus the MC operates under command feedrates. Therefore, the feed time should be calculated after giving careful considerations to the accelerations.

2.3 NC program evaluator programming

The NC program evaluator was programmed in Visual C++ under Windows 95. It is object-oriented and has a friendly user's interface. The NC codes which can be analyzed by the evaluator, besides such basic codes as G00-G03, include M98 (sub-program), G65 and G66 (macro), and optional code G202 (high speed mode for specifying feed accelerations). From the viewpoint of evaluating NC programs, the evaluator can be executed in the modes of single or consecutive blocks. Besides, the plane, origin and scale, *etc.* for tool paths, can be chosen optionally. The tool paths will be drawn on the screen when each block is interpreted while at the same time the operational times are calculated and saved into data files.

2.4 Analyses of sample NC programs

The NC program evaluator has been applied to analyze some NC programs. Table 1 lists the operational times with the comparisons between the actual and analyzed times. From the table, we can see that the estimated error is no more than 4% and the average rate of the real cutting time to the total time for the practical NC programs is 93.8%. It is much higher than that of NC lathes which was 82.8%².

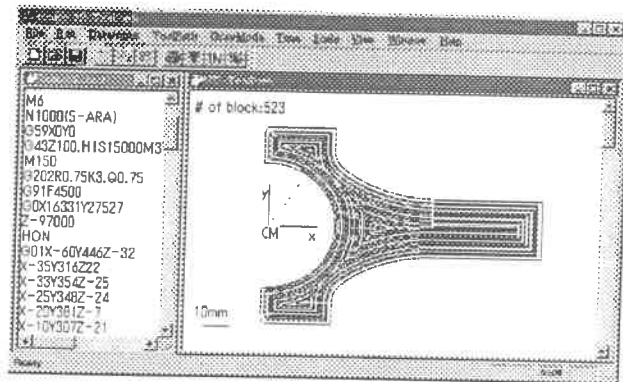
In the analyzed NC programs, Conrod is used to cut a forging mold. Its NC program contains G202 code and has 40,240 blocks. End-mills were used and changed four times. In Fig. 5, (a) shows the tool paths and (b) shows the ratios of the operational times.

Furthermore, the quantitative effects of feed rates and/or accelerations on the productivity can be analyzed by changing their values. For Conrod, the cutting time will decrease by only 9.8% when the feedrates is doubled. In contrast, it will increase by 50.5% when they are reduced by half. This is because there are a large number of tiny blocks and the feedrates are limited by G202 code in the NC program. The result implies that only increasing feedrates is not an efficient way to raise

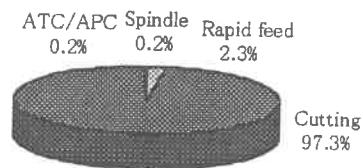
the productivity for this kind of NC program but decreasing feedrates will cause rather low productivity.

Table 1 Analyzed results of NC programs (min)

Name	Rapid feed time	Cut. time	Total time	Actual total t.	Real cut. time rate
Conrod	1.63	71.2	73.1	75.2	97.3%
Iron	2.11	17.6	19.6	20.4	90.3%
Test	0.06	0.86	0.97	0.98	88.5%



(a) Tool paths



(b) Cutting time ratios

Fig.5 Analyzed results of Conrod NC program

3. Conclusion

The NC program evaluator used to analyze NC programs and evaluate their productivity has been developed. With the evaluator, (1) the tool paths can be drawn on a microcomputer, and their quality can be evaluated and even the faults such as the interference can be found, and (2) the operational times and the cutting conditions can be evaluated.

In conclusion, the productivity can be evaluated by analyzing NC programs and improved by modifying NC codes associated with the cutting conditions and the parameters associated with the MC.

References

- 1 Huang S.H. and Zhang H.C. On the Use of Knowledge-Based Connectionist Models in CAPP Systems. *Manufact Sci & Engg*, ASME, 1994,1, 65-74
- 2 Shirase K., Hirao M., Yasui T. *et al.* Improving Productivity in NC Lathe Turning Based on NC Program Diagnosis System (1st report). *J Jap Soc Prec Engg*, 1994,60(2),235-239 (in Japanese)

LOCAL LAPPING SYSTEM TO OBTAIN WIDE AND FLAT SURFACE WITH ACTIVELY CONTROLLED FORCE SENSOR

Kiyoshi MATSUMOTO, Masayuki NAKAO and Yotaro HATAMURA
Department of Engineering Synthesis, Faculty of Engineering, The University of Tokyo
7-3-1 Hongo Bunkyo-ku Tokyo 113, JAPAN

1. Introduction

Machining process to obtain a wide and flat surface in short time is highly required in many industrial fields. In the field of semi-conductor technology, enlargement of silicon wafers implies an increase in the number of chips available at one process, and results in a decrease in production cost. But obtaining large wafers of, for example, 12" diameter and 0.3 μ m flatness by whole lapping is inefficient from view points of process time and material cost. The authors make an attempt on local lapping /grinding, which enables to lap or grind a surface locally to follow a prescribed trajectory.

2. Local lapping/grinding system with active force sensor
To achieve local lapping/grinding, the following problems need to be addressed. Because of the insufficient rigidity of the machining system, the mechanical elements such as the tool, the workpiece and the structural frame are deformed by lapping/ grinding reactive force. Fig.1 shows schematic of lapping/grinding process. If the tool is controlled so as to follow a prescribed trajectory in an open-loop fashion with no consideration on such deformation, the resulting surface does not retain the desired accuracy due to the deformation. Employment of a force sensor to detect lapping force is effective to monitor the deformation, but the presence of the sensor itself may cause a de-

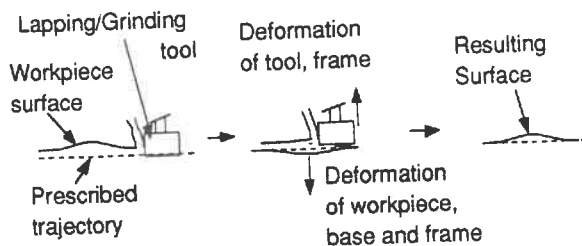


Fig.1 Schematic of lapping/grinding process

crease in rigidity. The authors previously proposed a force sensor with actively controlled compliance. 1) This sensor is actually equipped with an actuation capability of force or displacement as well as sensing capability in both. With this sensor, it becomes possible to control a tool position or lapping/grinding force while detecting the position and the force. A desired surface profile can be obtained in one pass by controlling this active tool with the force information and the knowledge about the rigidity of the lapping system. This operation consists of the following steps. First, the deformation of the sensor is compensated by actuating the device by the equal amount. Second, the deformation of the machining system can be canceled by detecting the reactive force and actuating the tool based on the knowledge about the rigidity.

3. Development of a ring shape active force sensor

3.1 Basic structure

Active force sensor utilizes the parallel plate structure. Fig.2 shows the parallel plate structure and its applications. To detect force, the basic parallel plate structure shown in Fig.2 (a) is used. Two pillars are connected by

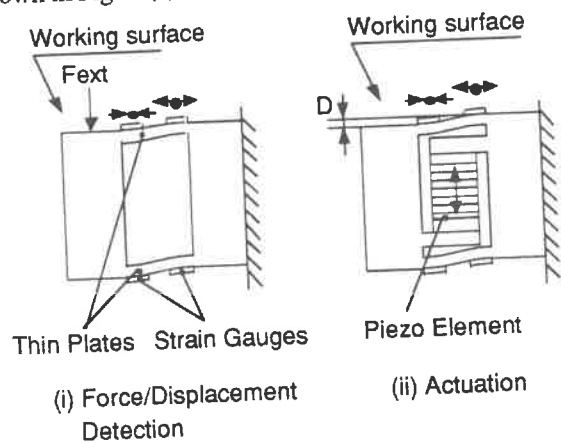


Fig.2 Two applications of parallel plate structure

a pair of parallel thin plates. One pillar is fixed to the rigid base, and another pillar is movable in one direction. The movable pillar has a working surface where external force or displacement is applied. On the surface of thin plates, strain gauges are glued. They detect deformation of the thin plates through surface strains, and thus applied external force or displacement can be detected. The structure shown in Fig.2 (b) is used for actuation. Between the pillars, a Piezo element is inserted. By elongation or contraction of the piezo element, displacement of the working surface is generated.

This Fig.3 shows the basic structure and the deformations of active force sensor.

A pillar is composed of upper/lower seats, and a piezo device which is glued between the seats. This pillar is movable and connected to a fixed pillar by a pair of parallel thin plates. The parallel thin plates limit movement of the movable pillar and movement of the piezo device in one direction. On the surface of the upper and lower thin plates, strain gauges are glued which detect their deformations respectively.

When the piezo element does not elongate, applied external force causes displacement of working surface D as shown in Fig3 (b) . Displacement D is detected by the upper and lower strain gauges and produce outputs E_u and E_l . When external force is not applied, elongation of the piezo element causes displacement of working surface D as shown in Fig (c) . Displacement D is detected by the strain gauges and produce outputs E_u and E_l similarly. The active force sensor uses superposition of these deformations. Assuming that the signs of E_u and E_l are positive as shown in the case of Fig.3 (c) . $E_l - E_u$ is proportional to the applied external force, and is used to detect applied force. E_u is proportional to the displacement of the working surface, and is used to control actuation by the sensor.

3.2 Ring shape active force sensor

Fig.4 shows structure of ring shape active force sensor. Between upper and lower ring, 4 active sensing units are installed at every 90 degrees. Each unit can detect force and actuate independently. By calculation of the outputs from 4 units, force in the direction of z and moments

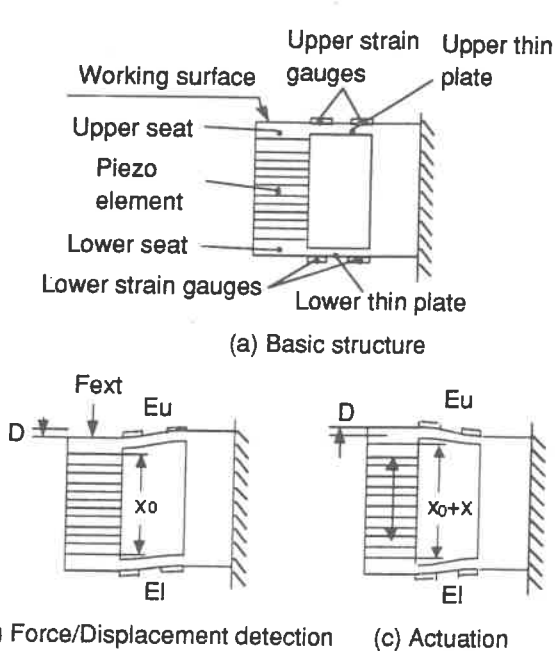


Fig.3 Basic structure and deformations of active force sensor

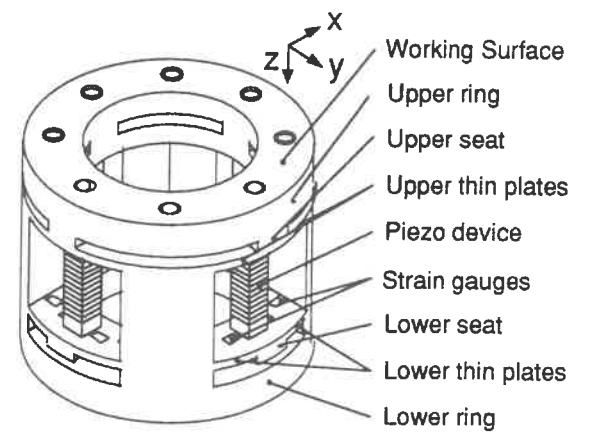


Fig.4 Ring shape active force sensor

around the axes x and y can be obtained. By actuation of the piezo elements in 4 units, displacement in the direction of z and inclinations around the axes x and y can be generated. The lower ring is fixed to frame or base, and machining tool is attached to the upper ring.

3.3 Sensing and actuating system

Fig.5 shows a block diagram of the sensing and actuating system. The output signals from 4 sensing units are expressed by suffix. E_{ui} indicates displacement of the work-

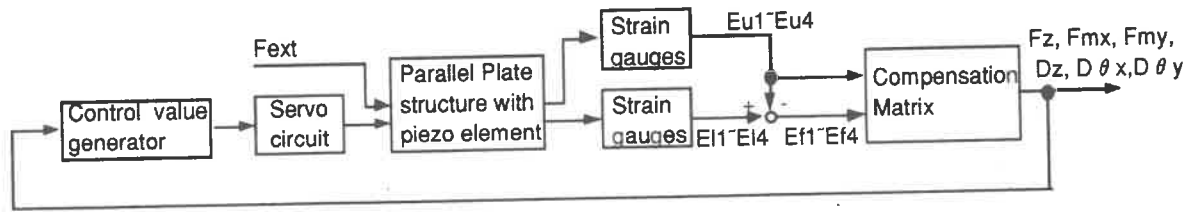


Fig.5 Block diagram of sensing and actuating system

ing surface detected by each unit. E_{fi} produced by subtraction E_{ui} from E_{li} indicates force detected by each unit. These signals have interferences from each other. Force and Displacement information is calculated by using compensation matrix to eliminate the interferences. As mentioned below, compensation matrix is obtained by calibration of the sensor.

Force and displacement information are fed back to control value generator. The control value is calculated so that the sensor acts as a spring which has arbitrary compliance. The control value drives piezo elements in sensing units through servo circuit.

3.4 Characteristics of fabricated sensor

Fig.6 shows a photograph of an actually fabricated sensor block.

The rated value and the resolution of force in direction of z are 100N and 0.1N, respectively. The actuation range and the resolution are 10 μ m and 0.01 μ m, respectively. The parallel plate structures were designed so that the resonance frequency may be as high as possible and the output signal of strain gauges may be as much as possible. The thickness of the thin plates is 1.0 mm, the length is 8.5mm, and the width is 12 mm. Rigidity to the force in direction z is 17.3×10^6 N/m when piezo elements are not driven.

Fig. 7 shows measured transfer function from voltage supplied to piezo elements to outputs E_u , which is the response characteristics of the working surface. The natural frequency is about 1.2 KHz.

Fig.8 shows compensation equation to obtain force information from outputs E_{fi} . The compensation matrix is obtained by calibration using weights.

Fig.9 shows force detection output when external force in direction of z is applied. The output is expressed by units of μ strain. It has good linearity. The sensitivity is 24.96 μ

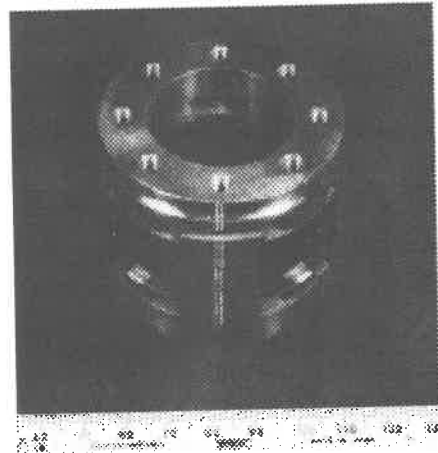


Fig.6 Photograph of sensor block

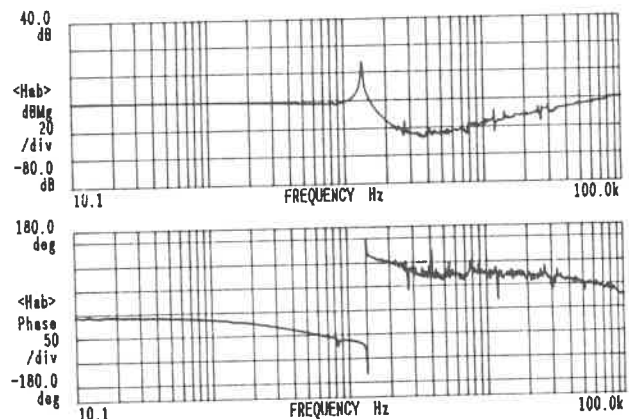


Fig.7 Transfer function

$$\begin{bmatrix} F_z \\ F_{mx} \\ F_{my} \end{bmatrix} = \begin{bmatrix} 2.00 \times 10^4 & 2.00 \times 10^4 & 2.10 \times 10^4 & 2.00 \times 10^4 \\ -9.43 & 4.92 \times 10^2 & -7.52 & -5.00 \times 10^2 \\ -5.11 \times 10^2 & -1.01 \times 10^5 & 5.07 \times 10^2 & -1.02 \times 10^5 \end{bmatrix} \cdot \begin{bmatrix} E_{f1} \\ E_{f2} \\ E_{f3} \\ E_{f4} \end{bmatrix}$$

Fig.8 Compensation equation for force detection

strain/N.

Fig.10 shows the force detection output to external force in direction of z when displacement of the working surface is changed. Displacement change has little effect on

force detection output. On the other hand there are upper and lower limits of displacement as shown by dotted lines. These come from the existence of the working range of the piezo element, and vary in proportional to applied external force.

Fig.11 shows the displacement of the working surface to applied external force. The compliance can be changed within the working range of the piezo elements.

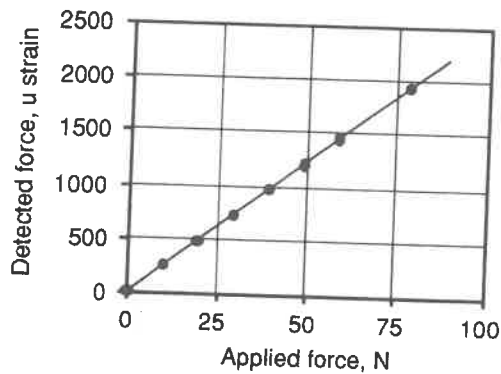


Fig.9 Force detection characteristics

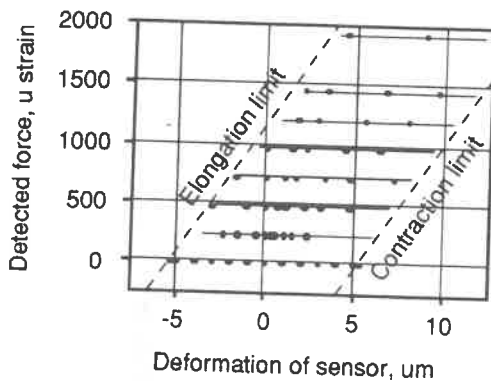


Fig.10 Force detection characteristics

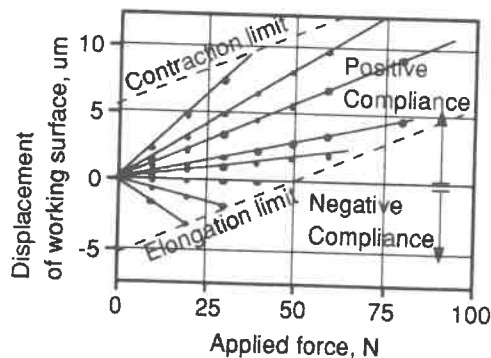


Fig.11 Compliance control operation

4. Experiment

4.1 Experimental system

Fig.12 shows the experimental system. A surface grinding machine is used as feed mechanism. Straight cup wheel is employed as grinding tool, and is rotated by a water driven turbine. The pressure of water fed to water inlet is about 3 MPa, and nominal rotation speed is 10000 r.p.m. Rotation shaft which connects the turbine and the grinding tool is suspended by a dynamic water pressure bearing. The ring shape active force sensor is installed between the turbine and the feed mechanism in z direction. An 8-inch silicon wafer is glued to the base by hot wax. The base is fed in x direction. The rigidity in direction z of this system without the sensor is measured to be 8.15×10^6 N/m, while that of the sensor is 17.3×10^6 N/m as mentioned before.

Fig.13 shows photograph of the experimental setup.

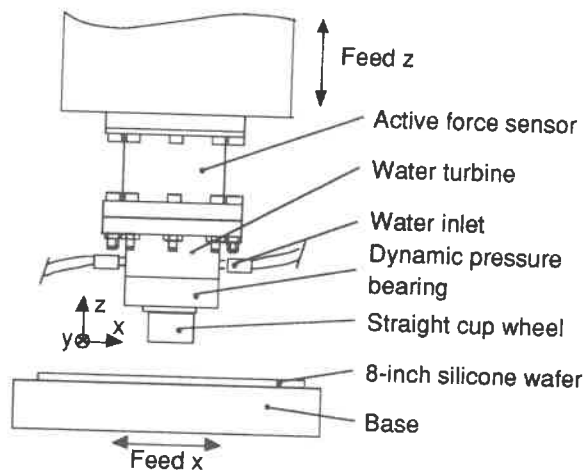


Fig.12 Experimental system

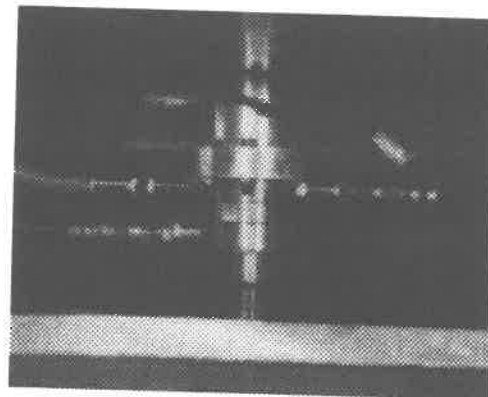


Fig.13 Photograph of experimental setup

4.2 Measurement of reactive force

Two straight cup wheels are used for experiment. The roughness numbers of abrasive grains on the wheels are 350 and 800. Feed rate is fixed at 200 mm/min. Reactive force of grinding measured is shown in Fig.14. Defining that the increase of the reactive force F_z to the depth of cut as dynamic rigidity K_d , K_d is $1.5 \times 10^5 \sim 2.1 \times 10^5$ N/m with the wheel of No.800. With the wheel of No.350, K_d is 8.5×10^4 N/m. Assuming that the rigidity of the grinding system as static rigidity K_s , K_s is 8.15×10^6 N/m as described in 4.1. K_d is $10 \sim 100$ times lower than K_s .

Fig.15 shows distribution of the reactive force at the depth of cut 50 μm with the wheel of No.800. Regarding that the feed direction as 0 o'clock and the tool rotation as CW, maximum force arises at 3 o'clock. The center of reactive force calculated from F_z , F_{mx} and F_{my} is around the inner cutting edge in direction of 4 o'clock. When the feed direction is opposite, the maximum force arises at 9 o'clock

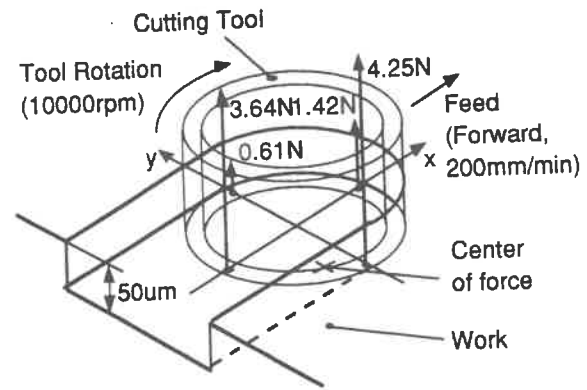


Fig.15 Distribution of reactive force

and the center is around the outer edge of 10 o'clock.

4.3 Grinding experiment with controlled compliance

In this experiment, reactive force was measured when grinding over a trapezoid shape of 100 μm , with variations of compliance of the sensor. The cup wheel is No. 350, and the feed rate is 200 mm/min. From the result of measurement in 4.2, the nominal reactive force is estimated to be about 8.5 N. Considering that the static rigidity of the grinding system K_s is 8.15×10^6 N/m, the total deformation including that of base, frame and tool is estimated to be about 1.04 μm , which may remain as grinding error. Bandwidth of the servo system to control the compliance of the sensor is set at about 10 Hz. DC gain is about 60 dB.

Fig. 16 shows experimental results when the rigidity of the active sensor is set at -19×10^6 N/m. The sign minus means that the compliance is negative, and that the sensor actuates in opposite direction of applied force. The height of the trapezoid is measured to be 93 μm . The reactive force during grinding is about 11 N, and the sensor generates deformation of 0.5 μm to cancel the relief of the work. The grinding error is measured to be 0.9 μm .

Fig.17 shows grinding error to compliance of the sensor. Compliance is inverse from rigidity. When the compliance is 0 N/m, which means the rigidity of the sensor is infinite or there is no sensor installed, the grinding error is 1.2 μm . Minimum grinding error 0.7 μm is obtained when the compliance is set at -0.053 N/m. The grinding error is reduced to 60%. In the region where the compliance is

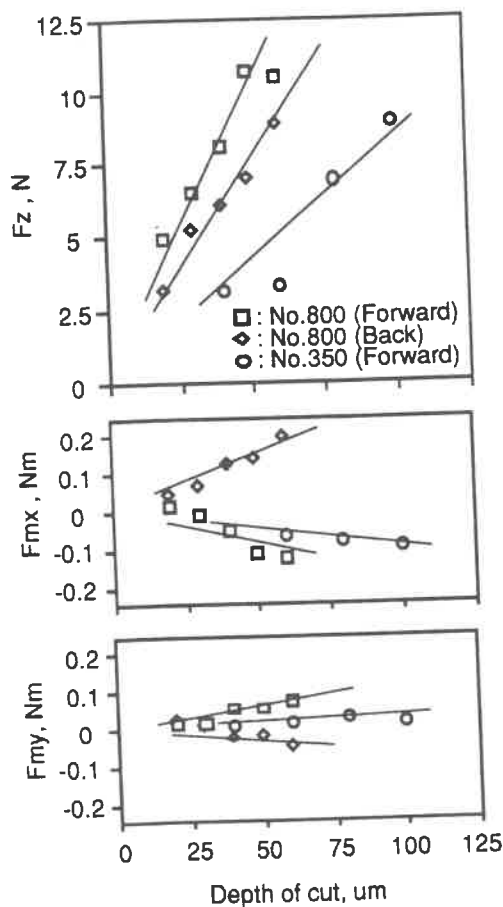


Fig.14 Reactive force of grinding

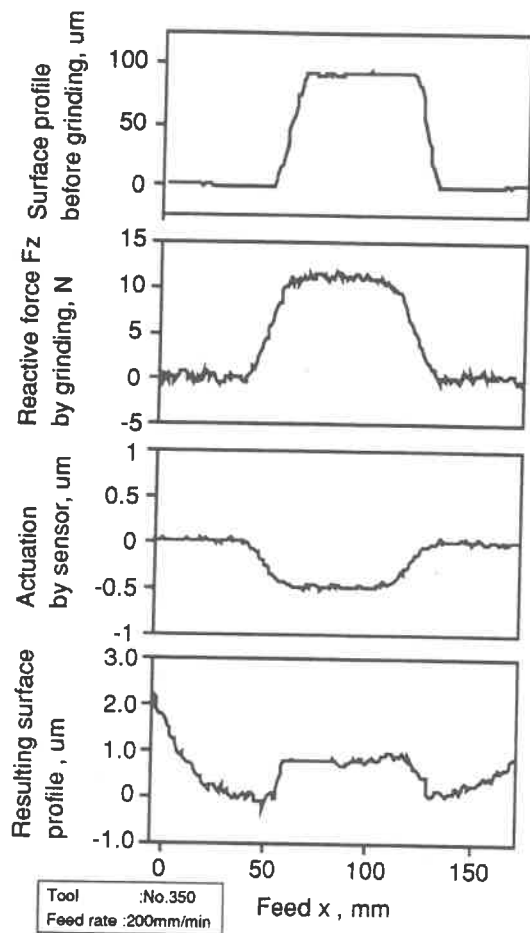


Fig.16 Grinding experiment
with negative compliance

less than -0.053 N/m , the servo goes unstable and the piezo elements elongate to their limits. The stability to the compliance has to be studied further.

5. Conclusions.

A local lapping / grinding system to obtain a wide and flat surface has been proposed. With this system, a significant decrease in process time and cost will be expected.

A ring shape active force sensor was developed which can change its compliance arbitrary.

The ring shape active force sensor was applied to surface grinding system with water driven turbine motor. In the experiment of grinding magnetic disk substrates, it was confirmed that grinding error was reduced to 60 %.

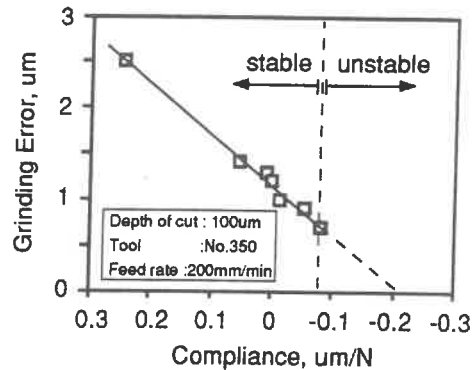


Fig.17 Grinding error to compliance

Acknowledgment

The authors would like to thank Mr. Masakazu Uesugi, Mr. Hiroyuki Ikeda and Mr. Katsutoshi Ohno of THK Co.Ltd. for discussion and supply of the water driven turbine motor and the dynamic water pressure bearing.

- 1) K. Matsumoto, et al., A Trial of Force Sensor with actively controlled Compliance, Proc. ASPE Spring Topical Meeting on Mechanisms and Controls for Ultraprecision Motion, ASPE, Tucson, USA, (1994)

Dynamic Friction Polishing of Diamond utilizing Sliding Wear by Rotating Metal Disc

K.Suzuki*, N.Yasunaga**, Y.Seki+, A.Ide*, N.Watanabe*, T.Uematsu++
 (*Nippon Inst. of Tech., **Tokai Univ., +De Beers Japan, ++Toyama Pref. Univ.)

Summary: Dynamic Friction Polishing Method which enables highly efficient abrasive-free polishing of a diamond workpiece is proposed. In the method a diamond workpiece is pressed onto a metal disc rotating at a high speed and polished by utilizing dynamic friction between them. It has been found that a stainless steel(SUS304) disc gives especially high polishing efficiency and that, on a square pillar shaped single crystal diamond(0.6x0.6mm with (100)plane), polishing efficiency of as much as $h'=2600\mu\text{m}/\text{min}$ ($0.94\text{mm}^3/\text{min}$) is attained under the conditions of sliding speed of $V_s=4000\text{m}/\text{min}$, stroke speed of $V_w=200\text{mm}/\text{min}$ and pressing force of $P=114\text{MPa}$.

Key words: Dynamic Friction Polishing, Abrasive-free polishing, Grinding of diamond, Stainless steel disc

1. Introduction

In general, diamond tools or diamond thin films are polished with a diamond wheel or a lapping plate. But there are some problems in the process in terms of machining efficiency and wheel wear. As one of the measures to solve these problems, a method utilizing thermochemical reaction between a heated ferrous metal disc and a diamond workpiece placed in a vacuum or H_2 chamber was proposed and good results were reported^{1,2)}. The authors have investigated another method utilizing dynamic friction between a diamond workpiece and a metal disc rotating at a high speed in the air, leading to abrasive-free polishing on a simpler device.

2. Dynamic Friction Polishing

In the method proposed here, a diamond workpiece is polished without using abrasives but by being pressed at a predetermined pressure onto a special metal disc rotating at a high speed to generate dynamic friction. The authors named this "Dynamic Friction Polishing (DFP) Method". Compared with the existing method utilizing thermochemical reaction, the new method has a bigger potential for practical application as it doesn't require a vacuum chamber or/and an equipment to heat a disc.

3. Experimental Equipments

A schematic illustration of the polishing equipment and its external appearance are shown in Fig.1 and Fig.2, respectively. A polishing tool was a metal disc with a diameter of 160mm (or 240mm in some cases) and was mounted on a spindle of a machining center(MC) with a tool holder. The lower surface of the metal disc was trued by cutting at a depth of cut of $a=0.03\text{mm}$ using a tool set on the machine table to minimize run out of the disc. A sample prepared as a workpiece was a square pillar shaped single crystal diamond with (100) planes

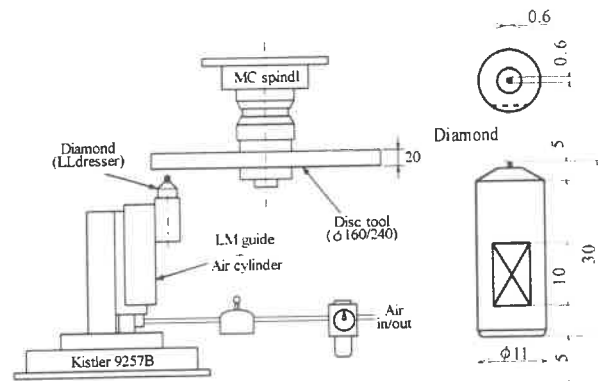


Fig.1 Schematic illustration of the Dynamic Friction Polishing and dimension of workpiece

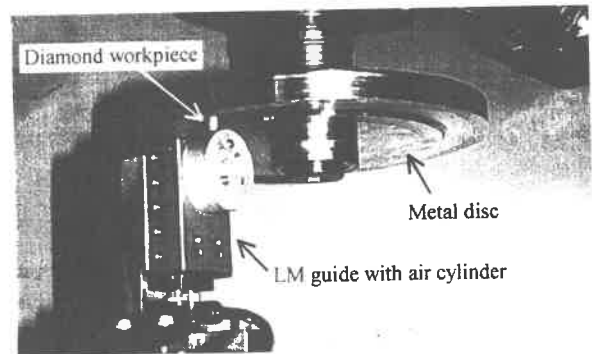


Fig.2 Abrasive free polishing experiment on MC

Table 1. Specifications of device and conditions

Device	Vertical MC, MT-V10(Mazak) + LM guide incorporated with an air cylinder
Disc	SUS304,S25C,S48C, $\phi 160\text{mm}(\phi 240\text{mm})$
Cutting condition	Cutting edge: SNGG120408 $V=96\text{m}/\text{min}, a=0.03\text{mm}, s=0.05, 0.1 - 0.5\text{mm}/\text{rev}$
Work-piece	Single crystal diamond Monodress($0.6 \times 0.6 \times 5\text{mm}$, upper face: 100), Holder: $\phi 11\text{mm}$ (De Beers)
Polishing condition	$V_s=1500, 2000, 2500, 2750, 4000\text{m}/\text{min}$ $V_w=10, 50, 100, 150, 200, 250, 300\text{mm}/\text{min}$, $P=50, 75, 100, 114, 130\text{MPa}$, $L=10\text{mm}$

on the both ends (0.6mm x 0.6mm, De Beers MONODRESS). The diamond sample was embedded in a column shaped holder and protruded from the holder by electric discharge dressing using a brush electrode. The protrusion height was limited to 700µm at the maximum to prevent breakage. The diamond sample was polished in a diagonal direction. Experiments were conducted in a dry condition.

4. Experimental Results

Conditions to achieve high efficiency in the Dynamic Friction Polishing method have been investigated.

4.1 Effect of a disc material

Polishing efficiency (wear rate of diamond) was investigated for various disc materials having different carbon contents (Fig.3). Pressing force (P) applied to the diamond sample was 100MPa, tool peripheral speed (Vs) was 2500m/min and a traverse oscillation (L) of 10mm (Vw=150mm/min) was given in a radial direction to ensure an even wear of the disc.

As shown in Fig.3, the less the carbon content of the disc material, the more the polishing efficiency increased and the disc of SUS304 showed particularly high efficiency. This result is thought to be attributed to the factors: 1) low carbon content (below 0.06%) giving advantages in carbon diffusion by thermochemical reaction, 2) low thermal conductivity (1/5 of iron) keeping the temperature at the top of the diamond sample without releasing the heat generated by dynamic friction. The fact that the lower the thermal conductivity of the rotating disc, the greater the temperature increase ($T-T_0$) at the point where dynamic friction is given can be explained also from the theoretical formula by Jaeger (Table 2). In addition, austenitic structure of SUS304 steel may also be advantageous to the diffusion of carbon.

As for the surface of the SUS304 disc used, only a slight crushing on top of the roughness profile was seen and no traces of being cut was observed (Fig.4). For

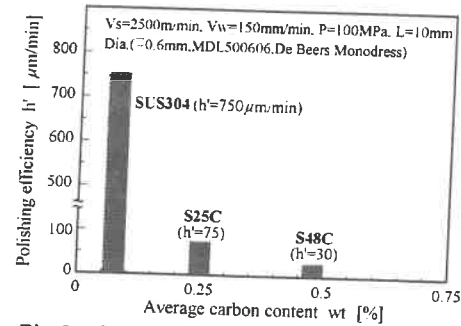


Fig.3 Disc material and wear rate of diamond

$$T - T_0 = \frac{x_1^{1/2} \cdot \mu \cdot W \cdot v}{3.76 \cdot l \cdot J \{ 1.125 k_2 \cdot x_1^{1/2} + k_1 \sqrt{l \cdot v} \}}$$

$$\text{here, } x_1 = k_1 / c_1 \rho_1$$

Table 2 Temperature rise at high sliding speed (Jaeger)

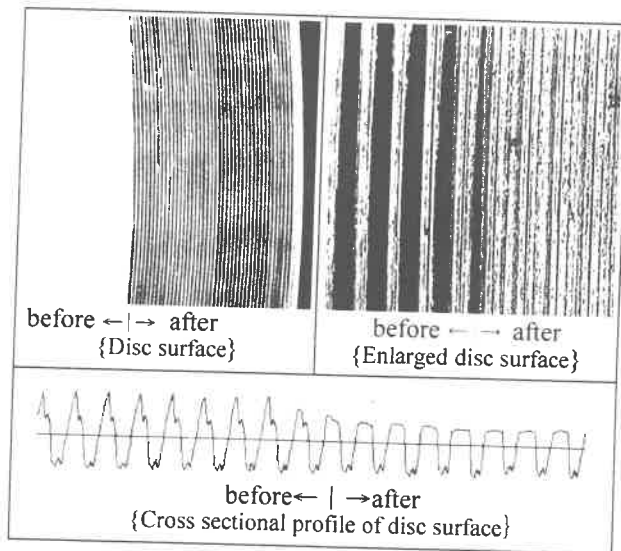


Fig.4 Surface and cross sectional profile of SUS304 disc

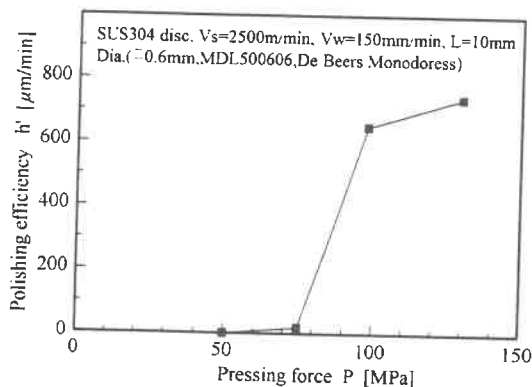


Fig.6 Effect of pressing force on wear rate ($V_s = 2500$ m/min)

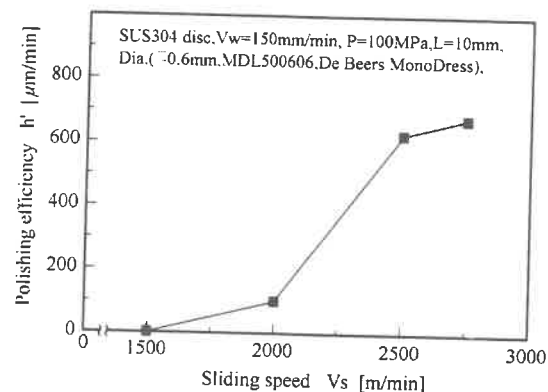


Fig.5 Effect of sliding speed on wear rate

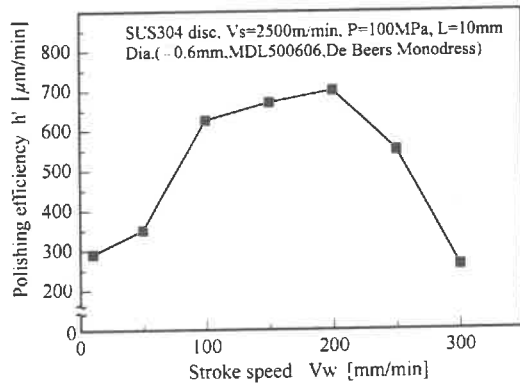


Fig.7 Effect of stroke speed on wear rate

reference, though no data is shown here, ceramics (alumina and zirconia) discs got worn without removing a diamond at all when tested under the same conditions.

4.2 Sliding conditions and polishing efficiency

The effects of polishing conditions were investigated using a SUS304 disc.

(a) Sliding speed: Fig.5 shows the effect of sliding speed on the polishing performance. When the speed was low such as below $V_s=1500\text{m/min}$, diamond couldn't be polished at all. From the fact that the polishing efficiency increases with the sliding speed in the speed range of 1500-2750m/min, far higher efficiency can be expected by further increasing the speed.

(b) Pressing force: Fig.6 shows the effect of pressing force on the polishing efficiency at the speed of $V_s=2500\text{m/min}$. Efficient polishing was performed when the pressing force greater than 75MPa was applied.

(c) Stroke speed: As shown in Fig.7, the maximum polishing efficiency was obtained when the stroke speed (V_w) was 200mm/min.

4.3 Roughness of the disc and polishing efficiency

By varying the feed rate of the cutting tool between the range of $s=0.1\text{--}0.5\text{mm/rev}$ when trueing the disc surface, the discs with different surface roughness were produced, and polishing tests were carried out with those discs using the conditions ($V_s=4000\text{m/min}$, $V_w=200\text{mm/min}$, $P=114\text{MPa}$) under which far higher efficiency could be expected. Roughness curves of the disc surfaces before and after polishing are shown in Fig.8 and the results are shown in Fig.9. In the range of surface roughness of $R_a=3\text{--}8.5\mu\text{m}$, no significant difference was found. But, in the range of $R_a=1.0\text{--}2.0\mu\text{m}$, polishing efficiency was quite high and extremely high value of $h'=2600\mu\text{m/min}$ ($0.94\text{mm}^3/\text{min}$) was achieved at the feed rate of $s=0.15\text{mm/rev}$ (Fig.10). This removal rate is more than 45 times of the figure

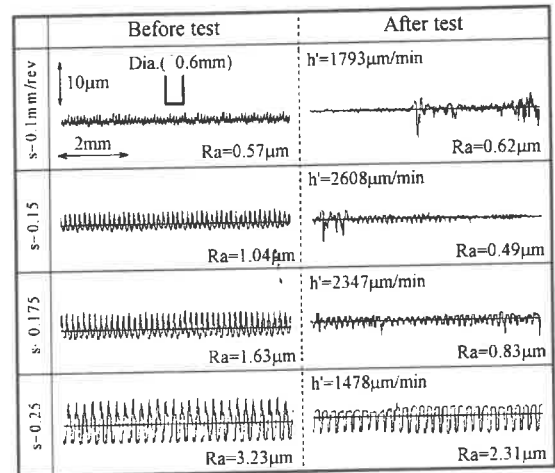


Fig.8 Surface roughness profile of sliding part on disc (SUS304 disc, $V_s=4000\text{m/min}$, $V_w=200\text{mm/min}$, $L=10\text{mm}$, $P=114\text{MPa}$)

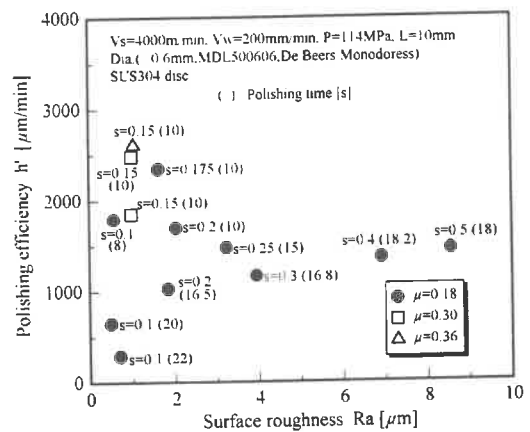


Fig.9 Relation between surface roughness of disc tool and wear rate

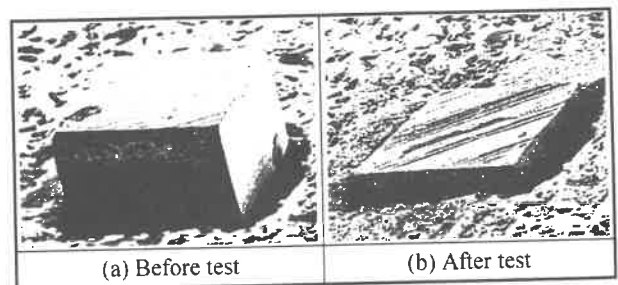


Fig.10 Rapid wear of diamond by DFP method (SUS304 disc, $V_s=4000\text{m/min}$, $V_w=200\text{mm/min}$, $L=10\text{mm}$, $P=114\text{MPa}$, $h'=2600\mu\text{m/min}=0.96\text{mm}^3/\text{min}$)

reported so far^{1,4)} as a record by abrasive-free methods. As shown in Fig.11, coefficients of friction in the cases of $h'=2450$ and $2600\mu\text{m/min}$ were rather high such as $\mu=0.3$ and 0.36 , respectively, compared with the typical figure of $\mu=0.18$. It is thought that higher friction heat generated by the increase of coefficient of friction would have resulted in the improvement of polishing efficiency (removal rate). The reason why such a

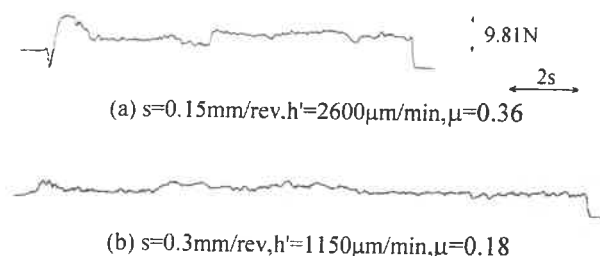


Fig.11 Change of tangential force at different disc surfaces (SUS304 disc, $V_s=4000\text{m/min}$, $V_w=200\text{mm/min}$, $L=10\text{mm}$, $P=114\text{MPa}$)

significantly high removal rate was achieved in the range of $R_a \approx 1.0\sim 2.0\mu\text{m}$ is not clear enough, but it can be thought that the peak shape of the roughness curve in this region was sharp enough to give rise to the increase of surface pressure at contact points, giving improvement in polishing efficiency.

4.4 Condition of the polished surface

Fig.12(a) shows the condition of the polished surface of the diamond when relatively high polishing efficiency ($h'=1300\mu\text{m/min}$) was obtained. It is seen that SUS304 material is adhered on the polished surface. As can be seen from the roughness curve shown in Fig.12(b), the surface roughness obtained by Dynamic Friction Polishing is not good enough. Though the roughness value might be improved by varying the polishing conditions, another method was taken here, i.e. SiO_2 mixed solution (Compol 40 by Fujimi Inc.) was sprayed on the polishing disc surface and further polishing for 4 minutes was given on the diamond. The result was that the adhered material was cleanly removed and the peak to peak value of the surface roughness was improved from $R_{\text{max}}=0.6\mu\text{m}$ to $R_{\text{max}}=0.06\mu\text{m}$ (Fig.13a,b).

Material removing mechanism in the Dynamic Friction Polishing method has not been made clear yet. But from the facts that even an inclined surface can be easily polished (Fig.14) and that sparks fly during the process, it is suggested that it might be due to transformation or diffusion of diamond by thermochemical reaction.

5. Conclusion

Using a disc of SUS304 rotating at a high speed, abrasive-free polishing was carried out on single crystal diamond workpieces and, by optimizing the conditions, high polishing efficiency such as $2600\mu\text{m/min}$ ($0.94\text{mm}^3/\text{min}$) could be achieved.

Moreover, it is expected that this method can be used for flattening a wide area of surface and also for grooving by just mouting a stainless steel disc with a suitable thickness on a surface grinder in the same way as in ordinary surface grinding.

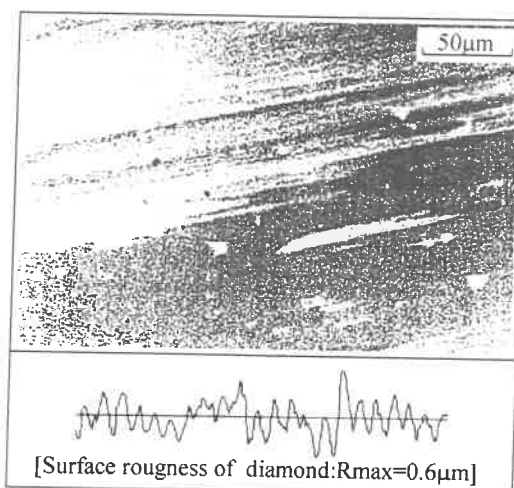


Fig.12 Condition of polished diamond surface by the Dynamic Friction Polishing method

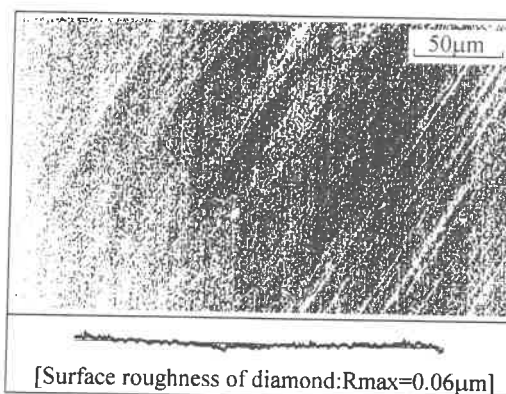


Fig.13 Improvement of polished surface by coroidal silica

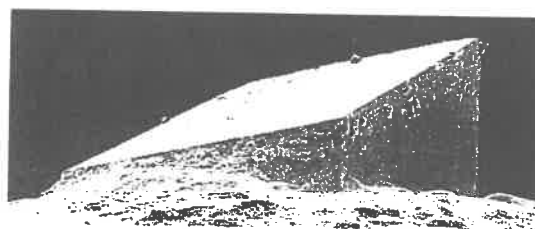


Fig.14 Slanted surface by the DFP method

In pursuing the study, the authors would like to express their heartfelt thanks to Noritake Diamond Co.Ltd., Ikegami Kanagata Kogyo K.K., Yamazaki Mazak Corp., Nisshin Seiko Co.Ltd., and Prof. S.Miyake of N.I.T. for their kind cooperation.

[References]

- 1) M.Yoshikawa et.al.: J.of JSPE 57-03(1991)106(in Japanese)
- 2) M.Yoshikawa et.al.: Proc. of JSDF's 8th Diamond Symposium (1994)16(in Japanese)
- 3) F.P. Bowden and H.G. Scott, Proc. of Royal Society, A248 (1958)368-378
- 4) N.Yasunaga et.al.: J.of JSPE, 38,4(1972)413(in Japanese)

Three-body Abrasive Wear Characteristics under Reciprocating Motion of CFRP in Vibrating Environment

Sadakazu TERAOKA *, Ken-ichi ISHIKAWA ** and Tatsuo NAKAGAWA **

* Kanazawa Technical College, ** Kanazawa Institute of Technology

* 2-270 Hisayasu Kanazawa Japan, ** Ohgigaoka Nonoich Kanazawa Japan

1. Introduction

Carbon fiber reinforced plastics (CFRP) has been widely used in industry because of their attractive mechanical characteristics.

Such CFRP parts are invariably subjected to three-body wear due to small indentations and machine vibrations.

In this study, the wear characteristics under the three-body condition and the abrasive wear of CFRP were investigated by using a vibrating environment and silicon carbide abrasive grains.

2. Experimental apparatus and method

The relation between the sliding direction and the lines of carbon fiber is defined as the rubbing direction. The L-direction is longitudinal and the T-direction is transverse, toward the lines of carbon fiber.

The abrasive wear apparatus is operated under the piston-reciprocated sliding friction based on the pin-on-plate method as shown in Fig. 1.

The CFRP test piece (length 20mm x thickness 5mm x width 5mm) ① is fixed to the sliding table ⑥ and it is rubbed on the plate (steel: S45C). The sliding table is moved by the crank ⑧ with the piston-reciprocated motion. The vibration action is created by using a mechanical oscillator ④. The test piece is loaded by the weight of the sliding table ⑥. The abrasive grain used for this

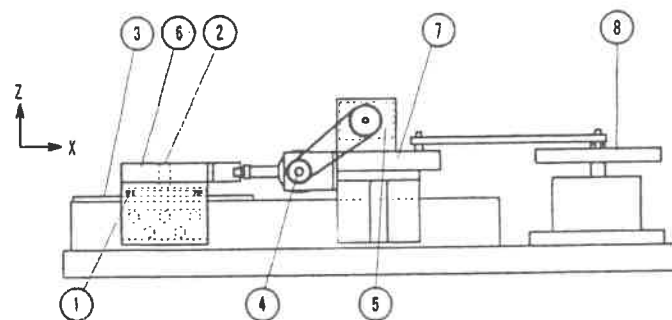


Fig. 1 Experimental apparatus

experiment is GC#600. A constant amount of the grain is supplied to the friction surface at hole ②, and it was discharged from the side of sliding table.

The wear volume is detected by the decrease an amount of the test piece thickness using the micro-meter. The vibratory direction is parallel to the sliding direction. The X-directional vibration is the same as the sliding direction.

The sliding speed was considered as an average speed of the crank. The sliding distance, when the vibration acted, is not considered such as the moving length of the frictional surface, but is dealt with as a real sliding distance (Eq.(1)).

$$Lx = \dot{X} \left[1 + \frac{2}{\pi} \left(\sqrt{V^2 - 1} - \cos^{-1} \frac{1}{V} \right) \right] \cdot t \cdots (1)$$

where, t : friction time(s), ω : circular frequency(rad/s), a : amplitude (mm), V : velocity ratio($a \omega / \dot{X}$)

Table 1 Experimental conditions

Sliding speed	$\dot{X}$	13.3mm/s
Frequency	f	0-20Hz
Pressure	P	0.39MPa
Amplitude	a	0, 0.3, 0.5, 1.0, 1.95mm
Abrasive particles		GC#600

The velocity ratio(V) was defined as the maximum sliding speed ($a\omega$) of the test piece per sliding speed($\dot{X}$) of the plate. When vibration acts, the over-lapping ratio is defined by Eq.(2), then the calculated wear rate is defined by Eq. (3).

Over-lapping ratio:

$$Ov = (\overline{V_X} - \dot{X}) / \overline{V_X} \quad \cdot \cdot \cdot (2)$$

Wear rate:

$$W = Wn \cdot (1 - Ov) \quad \cdot \cdot \cdot (3)$$

where:

Wn : wear rate at non vibration(mm²/m)

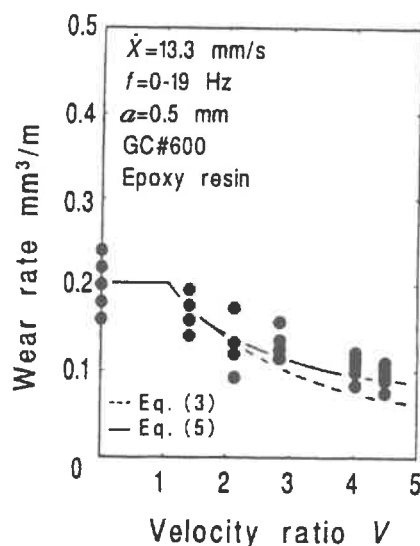
The experimental conditions are shown in Table 1.

3. Results and considerations

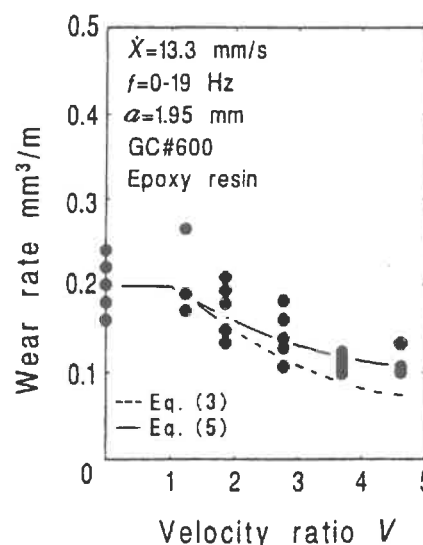
Fig.2 and Fig.3 show the relationship between the velocity ratio and the wear rate of CFRP and Epoxy resin. The dotted line shows the calculated wear rate, considering wear volume under the non vibration and the over-lapping ratio(Eq.(3)).

When the velocity ratio is from 0 to 1, the wear rate is subjected to non-vibration effects, because the vibration velocity of the test piece is smaller than that of the frictional surface. When the velocity ratio is more than unity, the wear rate decreases. The wear rate of Epoxy resin is smaller than the CFRP. When a flexural strength of the fiber is low, it is broken briefly at the transverse force and therefore, the effect of fiber reinforcement does not exhibit with the three-body abrasive wear.

In the amplitude 0.3mm (Fig. 3 (a)), the experimental value is agreement with the calculated value. In the amplitude 0.5mm (Fig. 3 (b)), the experimental results agree well with the calculated value as far as the velocity ratio 1.9. In the amplitude 1.0mm (Fig. 3 (c)), the velocity ratio is 1.4, and in the amplitude 1.95mm(Fig. 3(d)), its ratio is 1.2.

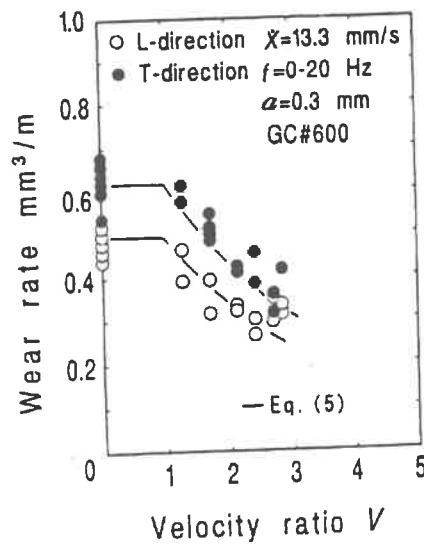


(a) $a = 0.5$ mm

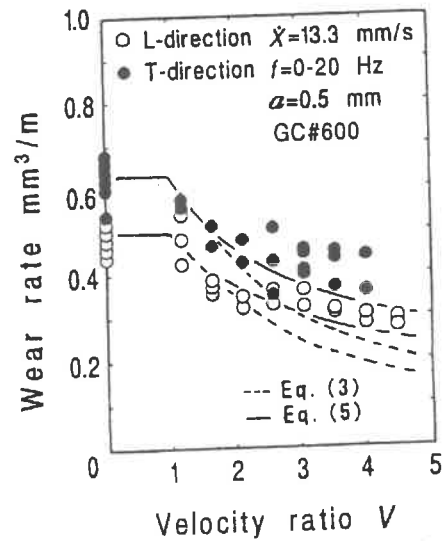


(b) $a = 1.95$ mm

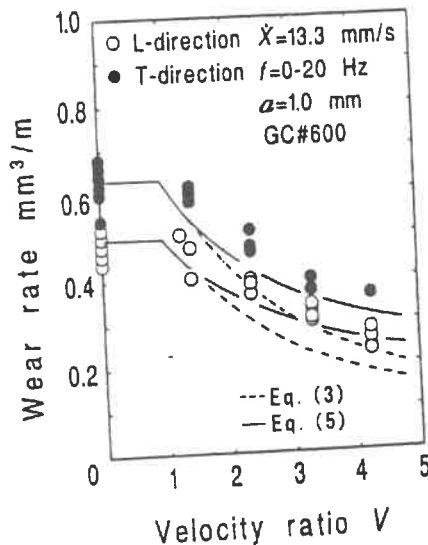
Fig. 2 Relationship between velocity ratio and wear rate (Epoxy resin)



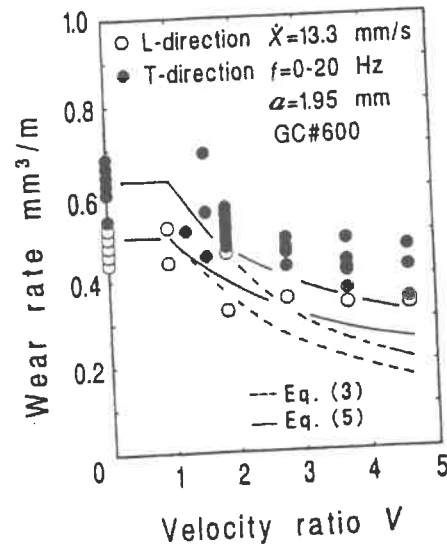
(a) $a = 0.3 \text{ mm}$



(b) $a = 0.5 \text{ mm}$



(c) $a = 1.0 \text{ mm}$



(d) $a = 1.95 \text{ mm}$

Fig. 3 Relationship between velocity ratio and wear rate (CFRP)

When the amplitude and the over-lapping distance increases, the motion of the abrasive grains become active rapidly. The over-lapping distance is defined by Eq. (4).

Over-lapping distance:

$$L_{OV} = 2 \cdot \left[a \left\{ \sin \left(\cos^{-1} \frac{1}{V} \right) \right\} - \frac{\dot{X}}{\omega} \cdot \left(\cos^{-1} \frac{1}{V} \right) \right] \quad \dots (4)$$

Fig. 4 shows the velocity ratio and the over-lapping distance. In any amplitude, the

over-lapping distance is 0.3mm, when the experimental value is not in comparison with calculate value.

The wear rate is calculated as follows:

- 1) In the velocity ratio from 0 to 1, the wear rate shows the same value at the non-vibration action.
- 2) In the over-lapping distance from 0 to 0.3mm, the wear rate was calculated by Eq.(4).
- 3) When the over-lapping distance is larger than 0.3mm, the place of the overlap is

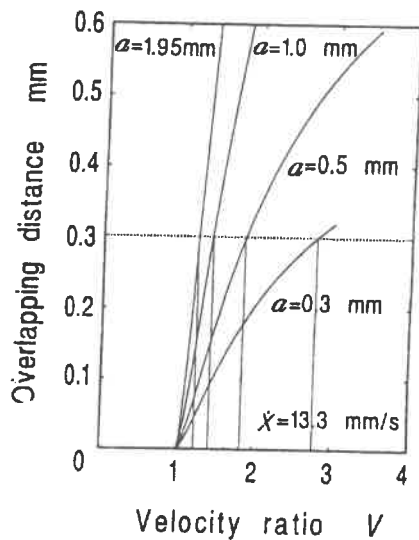


Fig. 4 Relationship between velocity ratio and over-lapping distance

influenced by the rolling and the scratching action of the abrasive grain. The value(Rov) is 30% mean frictional action. Then the wear rate is calculated by Eq. (5).

Wear rate:

$$W_{on} = W + W_n \left\{ 2(Lov - 0.3) / \bar{V}_x / f \right\} Rov \dots (5)$$

The calculated result is shown in Fig. 2 and Fig. 3 using a solid line. The experimental results agree well with the calculated value.

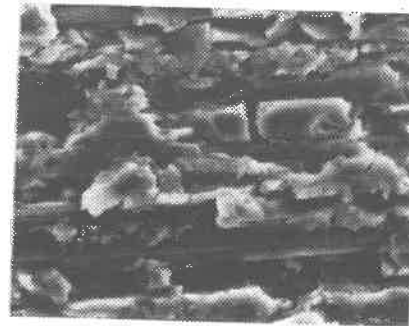
When the over-lapping distance is smaller than 0.3mm, the wear rate is not dependent upon the vibration. When the over-lapping distance is larger than 0.3mm, the wear rate depends on the vibration.

Fig. 5 shows observed results of CFRP by means of SEM. When the over-lapping distance is smaller than 0.3mm (Fig. 5 (a), (b)), the surface of CFRP showed a small indentation. As the over-lapping distance is larger than 0.3mm (Fig. 5 (c)), the breakage is sever.

4. Conclusions

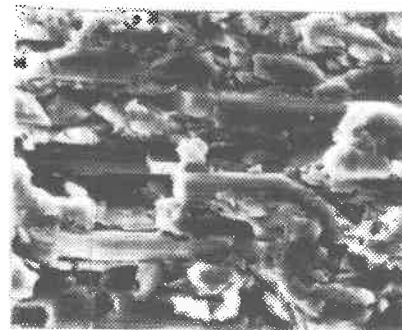
1) In the abrasive wear of CFRP, the effect of the fiber reinforcement is exhibited with

Sliding
direction
→

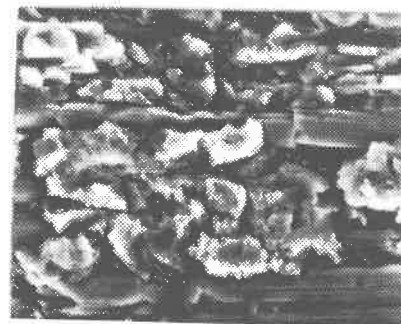


(a) $Lov=0mm$ ($V=0$)

Vibrational
direction
↔



(b) $Lov=0.24mm$ ($V=1.65$)



(c) $Lov=0.42mm$ ($V=2.36$)

High strength CFRP L-direction

$\dot{X}=13.3mm/s$ GC#600 $P=0.39MPa$ $a=0.5mm$

Fig. 5 Observation of CFRP

the two-body abrasive wear. However, its effect was not exhibited with the three-body abrasive wear due to the rolling action of the grains.

2) In the vibration action when the over-lapping distance is smaller than 0.3mm, the wear rate does not depend on the vibration.

3) When the over-lapping distance is larger than 0.3mm, the wear rate dose not depend on the vibration.

SURFACE INTEGRITY OF POLISHED INFRA-RED OPTICS AND GLASS

Prof. V.C. Venkatesh*, W.K. Chee*, Dr. K.L. Pey**, K.H. Tan* & Dr. A. Trigg**

* School of Mechanical & Production Engineering, Nanyang Technological University, Singapore 639798

** Failure Analysis & Reliability Department, Institute of Microelectronics, Singapore 0511

ABSTRACT

Polished brittle materials, both hard (glass, Si, and Ge) and soft (ZnSe, ZnS, and GaAs) were examined by SEM and AFM for residual scratches; the AFM has a better resolution and pictures of these are compared with SEM. Preston's coefficient for ZnSe and glass (plano, convex and concave) were determined and their values were found to closely match each other and those obtained by other international research workers. Polishing was found to be easier when ductile streaks were present when the partial ductile mode grinding process was employed. Single grit diamond machining was carried out on hard brittle and soft brittle materials and a distinct difference in pile up was observed.

1.0 INTRODUCTION

Brittle materials like Si, Ge, and glass are hard and brittle whereas ZnSe, ZnS, and GaAs are soft and brittle. The hard materials are easy to grind and polish whereas the soft materials are difficult to polish and need non-toxic chemicals [1]. In association with ophthalmic and infra-red optics industries, the polishing process was investigated and Preston's coefficients were determined [2, 3]. Polishing was found to be easier when ductile streaks were present when the partial ductile mode grinding process was employed [4]. Studies were carried out on the polished surfaces by SEM (scanning electron microscopy) and AFM (atomic force microscopy) to determine the extent of scratches left behind. Finally, single grit grinding was done on polished surfaces to study lateral displacement (pile up).

2.0 PRESTON'S COEFFICIENT

Preston's equation is a rate equation which simply states that the velocity of the surface recession, dT/dt , is related to the relative velocity between part and lap, ds/dt , through elasticity. This model assumed that a population of uniform size spheres penetrate the ideal plane of the workpiece in elastic Hertzian fashion, and in shear, plough a groove whose cross section is proportional to that of the penetration [5]. Preston's equation is defined as

$\frac{dT}{ds} = kp \frac{ds}{dt}$, where dT/dt is the layer removed per unit time, k Preston's coefficient, p the polishing pressure, and ds/dt is the relative speed between pad and substrate.

The results of polishing on glass and infra-optics are tabulated on Table 1. For ZnSe, Preston's coefficients are always the same order of 10^{-13} to 10^{-12} m²/N (10^{-14} to 10^{-13} cm²/dyne) [2]. Similarly, for glass, Preston's coefficients are of the order of 10^{-13} to 10^{-12} m²/N (10^{-14} to 10^{-13} cm²/dyne) [3]. These values coincide with the results of polished silicate and BK7 [5, 6]. Recent research has pointed out that Preston's coefficients for all kinds of standard glass polishing techniques are almost the same as long as a similar chemistry is applied [7]. The differences in material removal rate are merely due to different speeds and pressures.

Noting that Preston's equation describes the abrasive wear, the closeness of the results obtained seems to suggest that a similar set of abrasion phenomena underlay most polishing processes and that abrasion dominated the wear removal. However, it should be noted that only in gentle abrasion (which is the case for ZnSe and glass) plastic deformation dominates rather than hardness. The latter dominates in the more violent range of coarse abrasive removal [5]. At the same time, chemical influence cannot be ignored regardless of how minor the effect is. In fact, Brown, has indicated a strong chemical influence in the oxide abrasives CMP (chemical mechanical process) [5].

Table 1 Preston's coefficient for ZnSe and glass

Sample	Thickness removed during polishing, ΔT (μm)	Pressure, p (KPa)	Polishing time, dt (min)	Relative speed, ds/dt (m/s)	Preston coefficient, k (m^2/N)*
A (ZnSe)	78.4	137.9	240	0.1164	3.4×10^{-13}
B (ZnSe)	94	34.5	300	0.1164	1.3×10^{-12}
C (ZnSe)	41.3	13.4	300	0.1164	1.47×10^{-12}
D (ZnSe)	83.2	33.9	300	0.1164	1.17×10^{-12}
E (Plano Glass)	55	400	4	3.819	1.5×10^{-13}
F (Concave Glass)	161.5	400	1.75	3.496	11×10^{-13}
G (Convex Glass)	146	400	1.75	3.476	10×10^{-13}

* multiply m^2/N by 1×10^{-1} to give cm^2/dyne

3.0 SCANNING AND ATOMIC FORCE MICROSCOPY

The polishing process however good it may be, leaves behind marks. Polished surfaces when viewed under SEM barely reveal any marks. However, under AFM, polishing marks are so evident as to cause concern. However, infra-red optics materials had a quality similar to that of glass optical flats as far as surface roughness and flatness are concerned. Polishing marks of infra-red optics and glass imaged under SEM and AFM were compared. (Figure 1 - 2). The images obtained clearly depict the superiority of AFM in imaging the polishing marks over SEM. Table 2 gives a summary of the capabilities of the two equipment [8].

Table 2 Characteristics of SEM and AFM [8]

	SEM	AFM
Sample operating environment	vacuum	ambient liquid vacuum
Depth of field	large	medium
Depth of focus	small	small
Resolution		
x, y	5 nm	0.1 - 1.0 nm
z	< 1 nm	0.01 nm
Magnification range	10X - 10^6 X	5×10^2 X - 10^8 X
Sample preparation required	freeze drying, coating	none
Characteristics required of sample	surface must not build up charge and sample must be vacuum compatible	sample must not have excessive variations in surface height

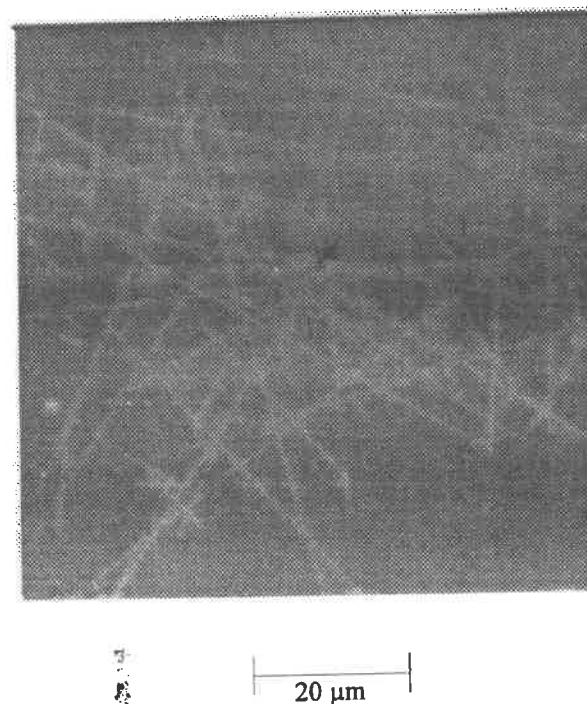
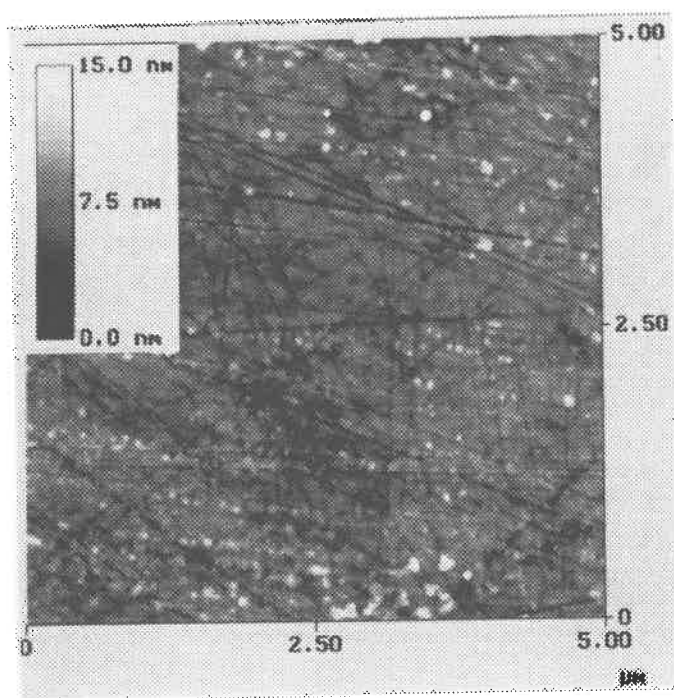


Figure 1 Surface topography of ZnSe imaged by AFM (left) and SEM (right)

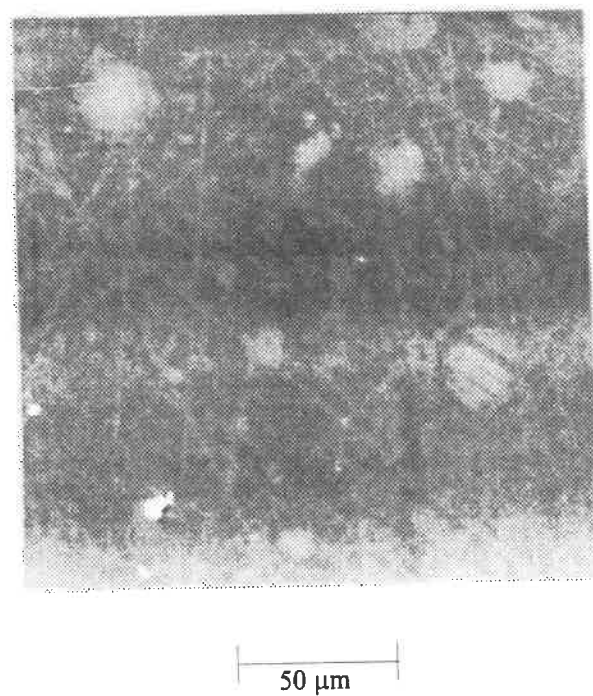
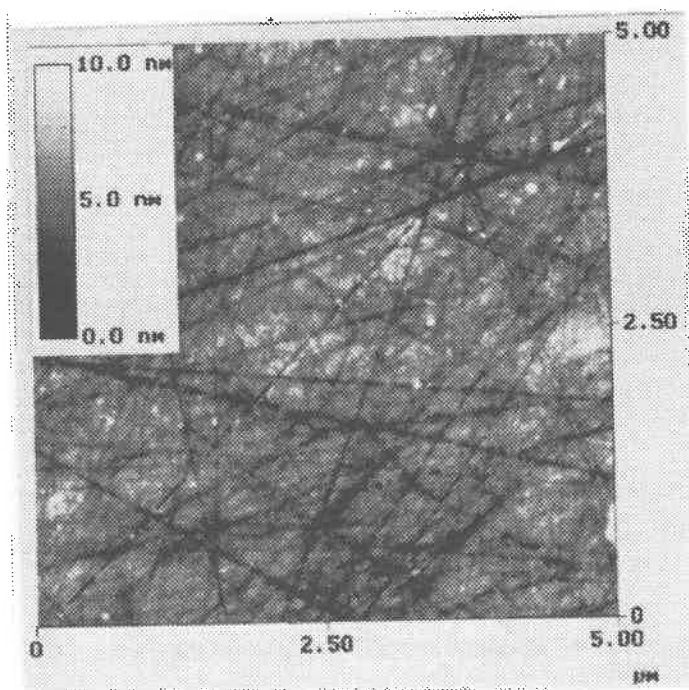


Figure 2 Surface topography of Ge imaged by AFM (left) and SEM (right)

4.0 SINGLE GRIT MACHINING

A conical diamond was used (Figure 3) for the single grit machining; given its accurate symmetry, the results of single and double pile up are due to material factors and not tool geometry. Single grit grinding was done on single crystal ZnSe (soft but brittle), Si, and glass (hard but brittle) using the Form Talysurf. It was observed that there is a distinct difference in the lateral displacement of the three materials. For ZnSe (Figure 4), equal pile up was observed, whereas only a single side pile up were observed on Si and glass (Figure 5-6). Reasons for this are still unclear and are being studied. The double pile up could be due to the softness and mono-crystalline nature of ZnSe. The single pile up could be due to the hardness and poly-crystalline nature of Si and glass.

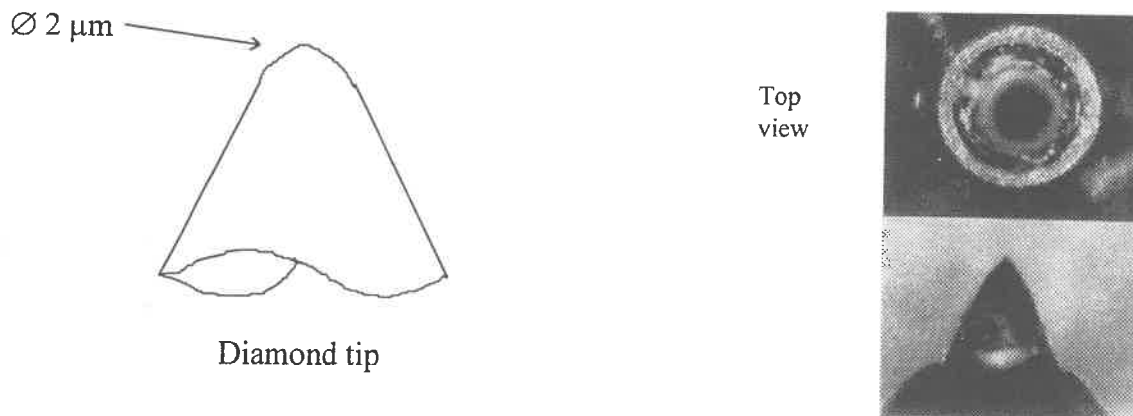


Figure 3 Coni-spherical diamond

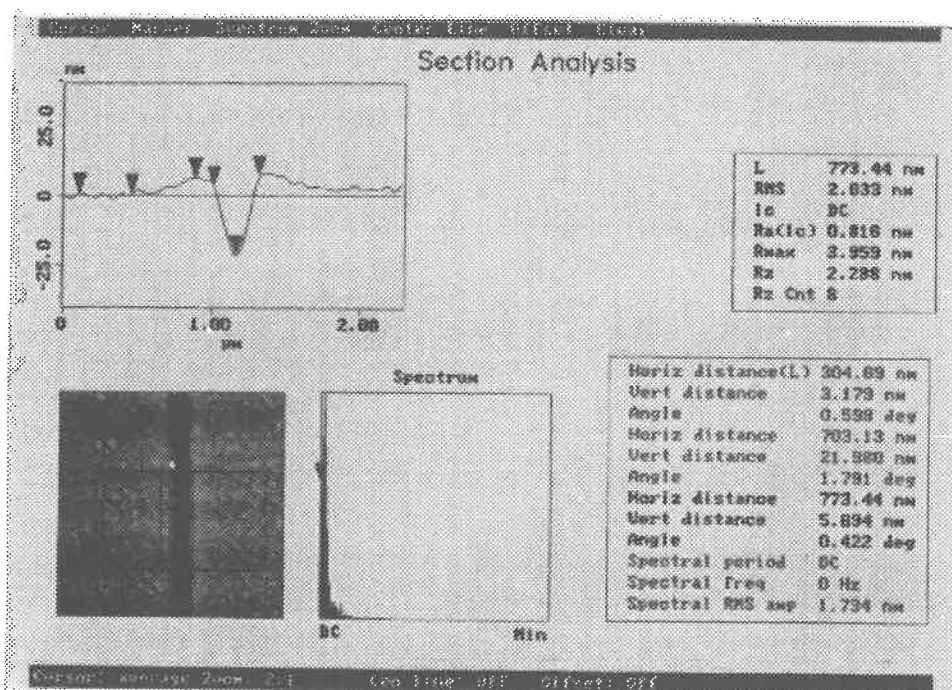


Figure 4 Ploughing action (without chip formation) causes equal pile up (ridges) on both sides for soft brittle mono-crystalline ZnSe

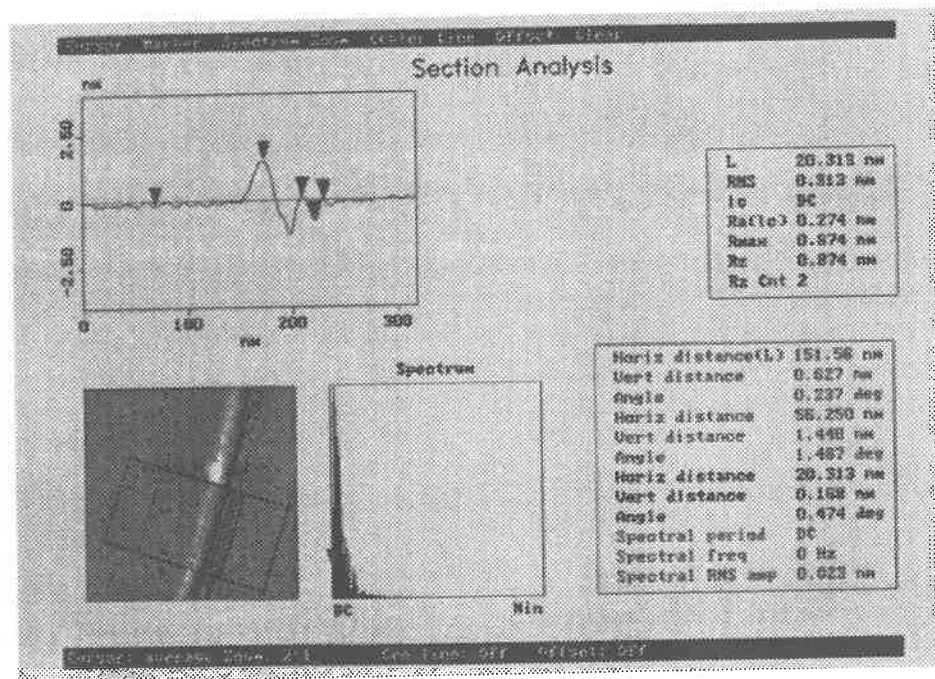


Figure 5 Pile up on only one side on hard brittle poly-crystalline Si

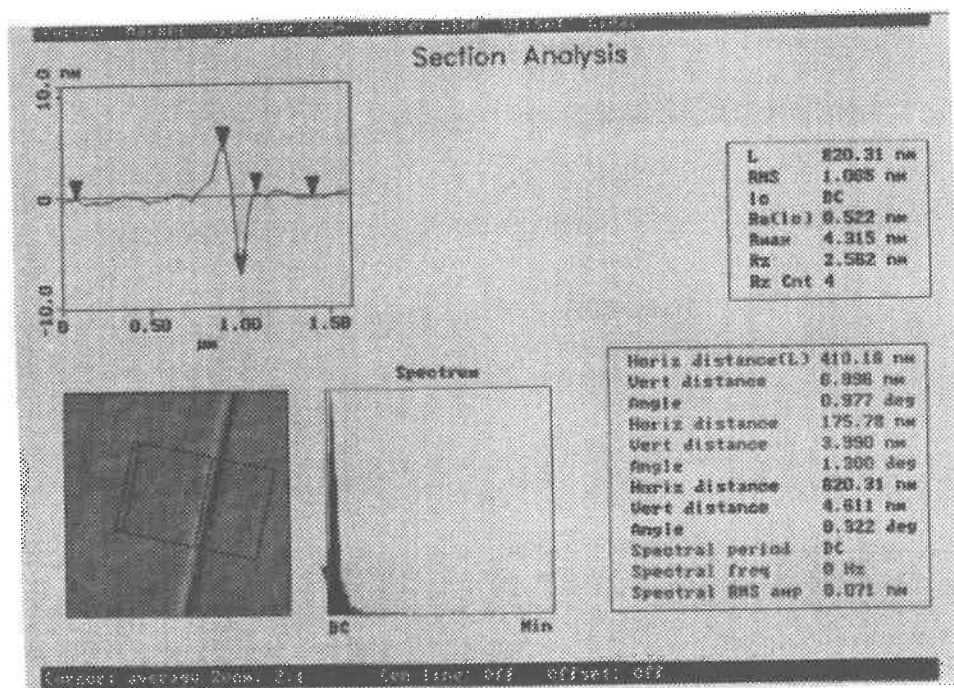


Figure 6 Pile up on only one side on hard brittle glass which is amorphous

5.0 CONCLUSIONS

1. The Preston's coefficient found for ZnSe and glass have the same order of magnitude (10^{-13} to 10^{-12} m²/N). This suggests that a similar set of abrasion phenomena underlay most polishing process {Table 1}.
2. The comparison of SEM and AFM shows the superiority of AFM in imaging fine polishing marks (Figure 1-2).
3. Double side pile up was observed on soft brittle ZnSe while single side pile up was observed on hard brittle glass and Si (Figure 4-6). Reasons for this are still unclear and are being studied. Possible reasons could be softness/hardness and mono/poly-crystalline nature of materials.

6.0 REFERENCES

1. Venkatesh, V.C., Inasaki, I., Toenshof, H.K., Nakagawa, T., and Marinescu, I.D. 1995. Observation on polishing and ultraprecision machining of semiconductor substrate materials. Annals of the CIRP 44/2: 611-618.
2. Chee, W.K. 1996. Surface integrity studies on hard and soft infra-red plano and spherical lenses. B.Eng. (Mech.) Thesis, Nanyang Technological University, Singapore.
3. Tan, K.H. 1996. A study of material removal during the polishing of precision spherical glass lenses. B.Eng. (Mech.) Thesis, Nanyang Technological University, Singapore.
4. Venkatesh, V.C., and Zhong, Z. 1995. Manufacture of spherical and aspheric surfaces on plastic, glass, and ceramics. Transactions of NAMRI/SME. Volume XXIII.: 669-174.
5. Brown, N.J. 1991. Lapping: polishing and shear mode grinding. Optical Society of America 3/1: 13-24.
6. Brown, N.J. 1987. Some speculations on the mechanisms of abrasive grinding and polishing. Journal of Precision Engineering 9/3: 129-138.
7. Fiedler, K.H. 1995. Processing (Grinding and Polishing). From Neuroth, B. (ed.), The properties of optical glass. Berlin Heidelberg, Germany: Springer-Verlag. 247-264.
8. Strausser, Y.E., and Heaton, M.G. 1994. Scanning Probe Microscopy Technology And Recent Innovations. Santa Barbara, USA: American Laboratory: 1-2.

Burr and shape distortion in micro-grooving of optical components

Jung-Hwan Ahn*, Han-Seok Lim*, David A. Dornfeld**

* Pusan National University, Pusan, KOREA

** UC Berkeley, CA, U.S.A.

1. Introduction

Micro-groove, ranging in size from a few microns to several tens of micron, is the basic geometric shape of miniaturized/thinned optical components such as fresnel lens, lenticular lens, reflector, etc. Requirements of high precision shape/surface quality for optical parts to function as designed need such a ultra-precision machining with a single point diamond tool. It has also a unique advantage that various grooves can be made only by changing the tool geometry or the machining program.

However, actual problems such as shape distortion and burr, which might be caused by high specific cutting force and ductile mold materials, is so undesirable because it absolutely degrade the performance of optical components. Many researches have been conducted about burr generation mechanism mostly in conventional machining[1,2]. Burr in a single groove, which swells up at both sides of a groove, was investigated.[3,4]

In this study, a simplified model of the burr/shape distortion in grooving multiple micro-grooves is proposed and the effects of cutting parameters such as depth of cut, groove angle, etc. are experimentally investigated.

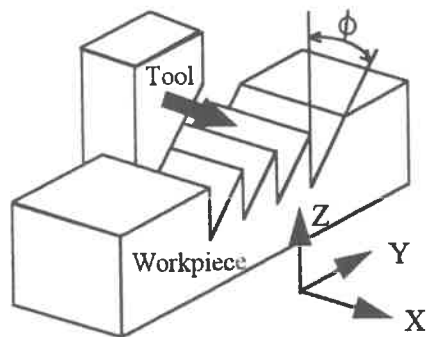
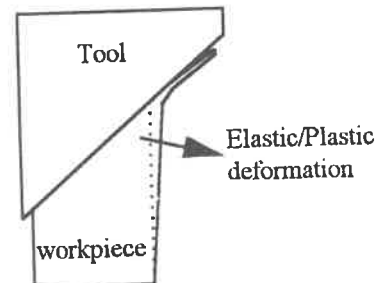


Fig.1 Array of micro-grooves and tool

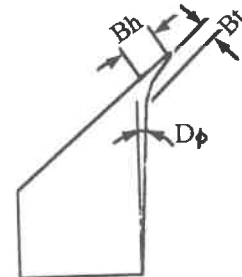
2. Model of burr and shape distortion in micro-grooving

The term "burr", cited here, is limited to the burr which is generated at the summit of groove tooth, and, thus is equivalent to side burr in milling and grinding. Fig.1 shows the geometrical shape of an array of micro-grooves and a tool. Fig.2(a) shows a simplified model of burr generation and shape distortion and Fig.2(b) shows designations for burr height(Bh), burr thickness (Bt) and shape distortion($D\phi$).

When the tool begins cutting, the groove tooth is elastically/plastically deformed by the thrust force normal to groove surface during chip separation. The pre-generated burr hanged on the tooth edge rotates away the cutting edge so that only a part of burr is cut away by the tool. The rest of the burr remains accumulated on the newly-generated burr. This is the formation process for burrs in multi-



(a) Deformation of workpiece



(b) Designation of burr/shape distortion

Fig.2 A model of burr/shape distortion in grooving

groove operations. That is, a combination of bending, shear and Poisson deformation due to both thrust force and the geometrically-weakened-tooth is assumed to cause the burr and shape distortion in multi-grooving.[5]

3. Experimental method

Fig.3 shows the schematic diagram of experimental setup. The self-developed machine has a microstepping motor driven z-axis with a resolution of 0.2 m and an AC servo motor driven XY table with a resolution of 1 m and is mounted on a granite plate with vibration-absorbed air springs. Cutting forces are measured with a tool dynamometer and recorded while grooving for afterward data analysis.

Cutting parameters to be investigated are depth of cut, groove angle, and workpiece material. Cutting speed, cutting edge radius, rake angle and tool setting angle are to be kept constant 5mm/s, 0.8 μ m, 0 deg and 0 deg respectively. The experiments were conducted

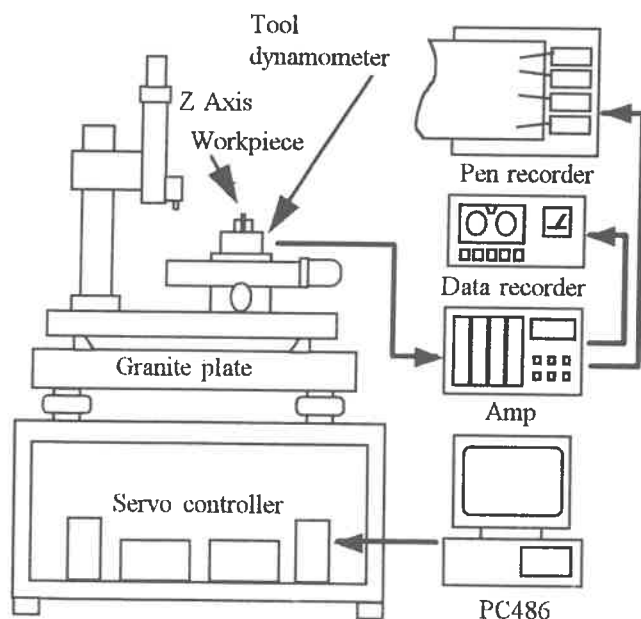


Fig.3 Schematic diagram of experimental setup

Table 1 Experimental cutting condition

scheme	speed	pitch	groove angle	depth of cut	work piece
single	5mm/s		20°-80°	1-12 μ m	brass Al7075 OFHC
multiple		50 μ m	50°	1-50 μ m	

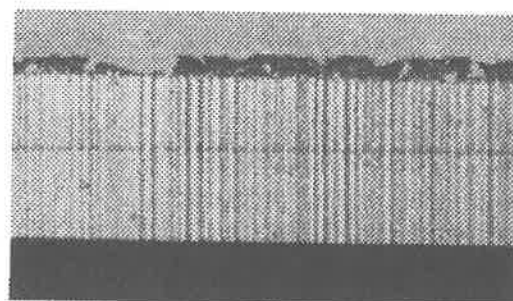
in 2 ways. One is, for the convenience of measuring cutting force and burr shape, the orthogonal cutting with 0.5mm thick workpiece, so called single-grooving, and the other is multi-grooving to measure the shape distortion. Every experiment under a given condition is conducted repeatedly up to the total accumulated depth of 100 μ m and the total amount of burr and shape distortion is compared. Cutting condition is summarized as Table 1.

4. Experimental results and discussion

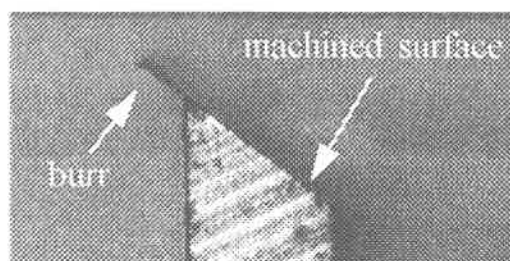
4.1 Burr shape and shape distortion

Fig.4 shows (a)top view and (b)side view of a machined surface. Burr develops outwards at the sharper side of groove surface while nothing happens at the blunt side. This implies that burr of micro-grooving is generated in other way than in milling or grinding.

Fig.5 is a microscopic photograph showing cross-sectional view of distorted gooves. Grooves are deformed opposite to tool feeding direction, because the stiffness of groove tooth is not enough to sustain the thrust cutting force. Proper cutting condition, therefore, should be selected to keep burr/shape distortion within a tolerance level.



(a) Top view



(b) Side view

Fig.4 Machined surface and burr

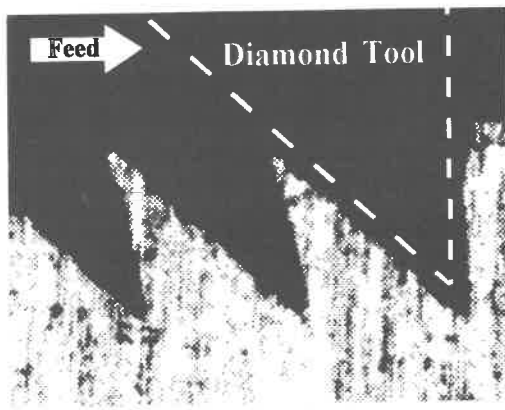


Fig.5 Shape-distorted groove

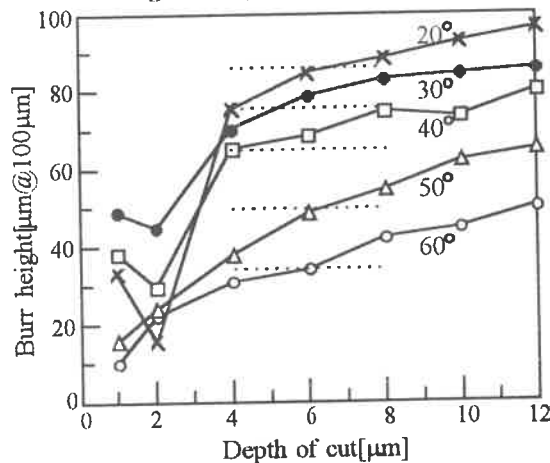


Fig.6 Accumulated burr height

4.2 Effects of depth of cut and groove angle

Fig.6 shows the accumulated burr height for 100μm depth-groove with depth of cut and groove angle varied. As depth of cut increases, burr gets thicker and larger. Burr height is nearly proportional to depth of cut.

As shown in Fig.7, in larger than 6μm of depth of cut, burr gets uniform. Meanwhile, in smaller depth of cut, burr becomes thinner and nonuniform with parts of burr torn out probably by the shear force and even small cracks in the boundary between burr and workpiece.

Smaller groove angle results in larger burr height. As the groove angle is getting smaller, the stiffness of tooth is getting lower and it causes the larger bending deformation which results in a larger burr. This demonstrates well the suggested model of burr in grooving.

As shown in Fig.6, for brass, 6μm may be a critical depth of cut where burr characteristics in view of uniformity and possibility of

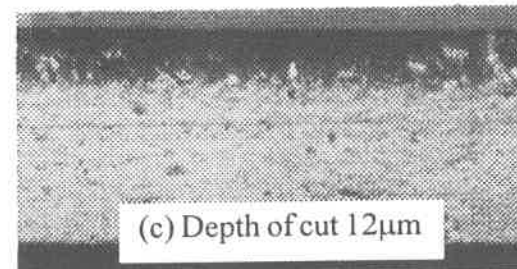
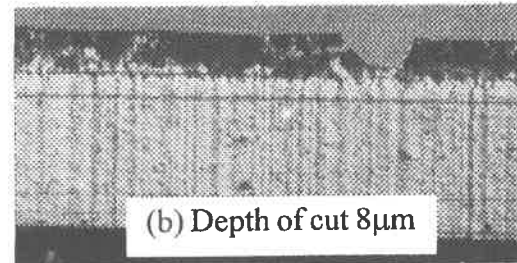
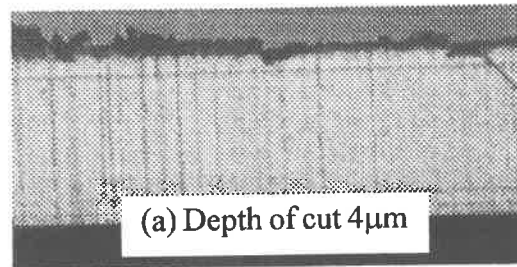


Fig.7 Comparison of burr shape

removal changes rapidly. The accumulated burr height at the depth of cut of 6μm is nearly equivalent to the projection(B_p) of the accumulated depth of 100μm on the machined surface. Dotted line indicates B_p for each groove angle. Cutting conditions to result in burr less than B_p would be recommended to help deburring following grooving.

Fig.8 shows cutting forces with depth of cut and groove angle varied. As depth of cut increases, both cutting forces increase almost linearly. As groove angle decreases, the thrust force increases largely whereas the principal force increases just a little. What is most important is that F_y , the y-directional component of thrust force, increases rapidly with groove angle decrease, as shown in Fig.8. It is considered due to the increase of elastic force reacted by the groove tooth which is largely elastically deformed. This also verifies why the burr is dominant at the sharp side in single-grooving.

4.3 Effects of workpiece material

Fig.9 shows the effect of workpiece

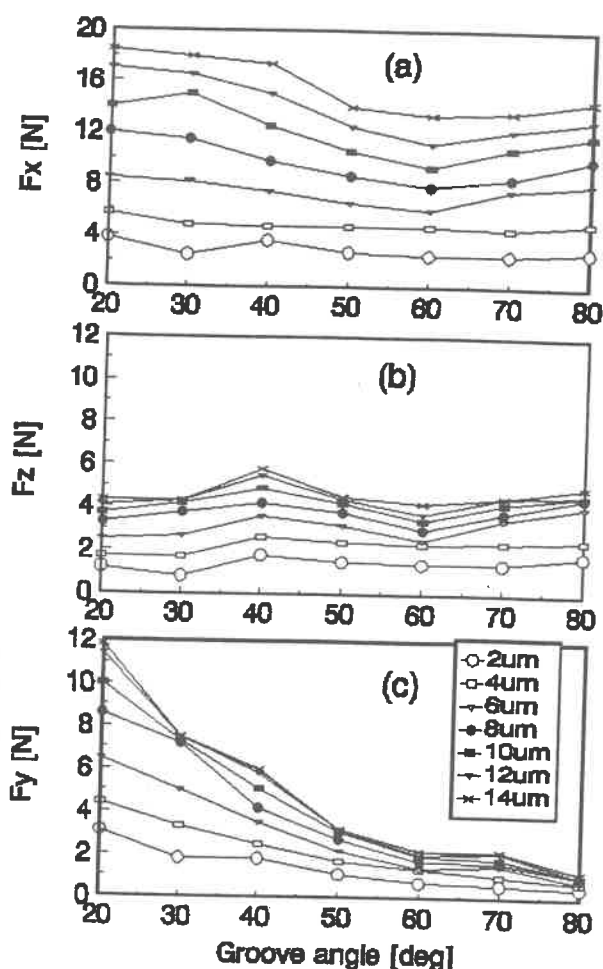


Fig.8 Variations of cutting forces
(a) principal force
(b) z-component of thrust force
(c) y-component of thrust force.

material to the burr height(a) and shape distortion(b) in multi-grooving. In Fig.9(a), the burr height increase with depth of cut like in Fig.6. Aluminium and copper seem to generate larger burr than brass because they are more ductile. As shown in Fig.9(b), shape distortion does not occur when depth of cut is less than 30μm because the strength of tooth can sustain the thrust force. Over 30μm shape distortion occurs nearly proportionally to depth of cut. Shape is less distorted in brass than in the others because the strength of brass is higher.

Therefore shape distortion is less significant once depth of cut is selected referring to B_p .

5. Conclusion

- 1) The thrust force in micro-grooving plays a

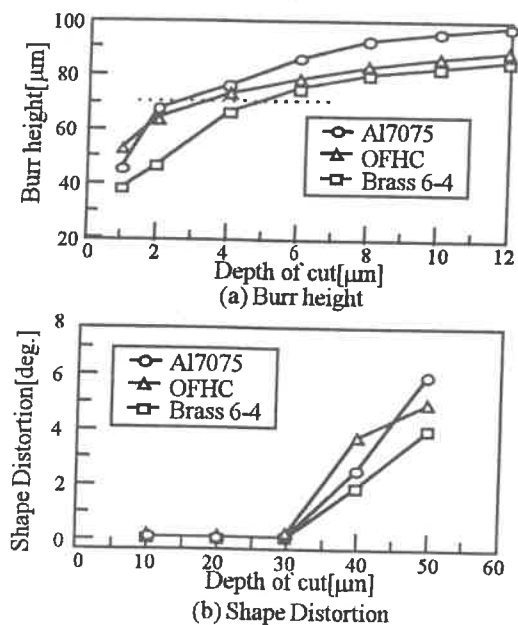


Fig.9 Accumulated Burr height and shape distortion of multi-grooving

significant role in burr generation and shape distortion.

- 2) Burr of micro-groove grows largely due to bending deformation caused by large groove angle and large depth of cut.

- 3) A critical depth of cut, where an accumulated burr height is less than the projection(B_p) of the accumulated depth of cut, is recommended for deburring after grooving.

- 4) Shape distortion is less significant when depth of cut is selected considering less burr.

References

- [1] L.K. Gillespie, P.T. Blotter, The Formation and Properties of Machining Burrs, J. of Eng. for Ind., Trans. of ASME, Feb., pp66-74, 1976
- [2] Sung-Lim Ko, David A. Dornfeld, A study on Burr Formation Mechanism, J. of Eng. Mat. and Tech., Trans. of ASME, Vol.113, Jan., pp75-87, 1991
- [3] Ueda. K., Sumiya. M., Ultra High Precise Micro Grooving, J. of JSPE, Vol.52, No.12, pp. 22-25, 1986(in Japanese)
- [4] Toshio Tamura, Noriyoshi Arakawa, Masami Masuda, Generation Phenomena of Burrs in Micro Grooving of Single Crystal Ferrite, J. of JSPE, Vol.57, No.5, pp.155-160, 1991.(in Japanese)
- [5] Kazuo Nakayama, Minoru Arai, Burr formation in metal cutting, Annals of the CIRP, vol.36, No.1, pp33-36, 1987

AN EMPIRICAL TOOL FORCE MODEL FOR PRECISION MACHINING

Christopher Arcona¹ and Thomas A. Dow²

INTRODUCTION

The components of the steady-state cutting forces in metal machining operations contain a wealth of information on the material removal process. Changes in these forces indicate changes in machining parameters, such as the depth of cut and tool sharpness. Thus, it should be possible to exploit tool force as a monitor of precision machining processes. The effective use of tool force in a feedback control scheme would require an understanding of the parameters that affect the force components and a model of the relationships between them. Unfortunately, an empirical model that would facilitate rapid data processing is not currently available for precision metal machining, although considerable effort has been expended in attempts to explain the cutting process [1-5].

A model of the machining process that is based on experimentally observed material flow is discussed in this paper. Chip formation is modeled as a process of shear where the angle of the shear zone is determined by the coefficient of friction between the chip and the tool. This proposed model also accounts for the contribution of elastic deformation of the workpiece to the total force on the tool and allows machining forces to be calculated from basic material parameters. Metals having dissimilar properties and grain structure (electroless nickel, electroplated copper, 6061-T6 and 5086 aluminum alloys) were studied. All machining experiments were carried out with a diamond tool having a 5.08 mm nose radius, 0° rake angle and 8° clearance angle.

TOOL FORCE MODEL

The chip structure of the four materials has been studied with the aid of a technique that produces high resolution images of the chip flow [6]. Consider the plated copper chip cross section in Figure 1. This chip was created with the freshly lapped diamond tool at a cutting speed of 25 mm/min. The morphology of this type of chip has been discussed elsewhere [6-8]. For this study, the important chip characteristic is the "shear angle," or the angle that the thin, highly deformed zones make with the horizontal. For the plated copper chip shown in Figure 1, the shear angle, ϕ , is about 45°, as indicated in Figure 2. Figure 2 depicts the same cross section with the outline of the next segment overlayed. This segment would form if the tool could be replaced and moved about 500 nm to the right. The forces on this representative section are indicated in the diagram. The total force acting on the tool in the cutting direction is F_c . The force in this direction associated with elastic deflection of the workpiece, or material spring back, is $F_{celastic}$. The tool forces due to this elastic deformation of the workpiece have been shown to be significant in

¹ Staff Engineer, International Business Machines Corp.; San Jose, CA 95193

² Professor, Precision Engineering Center, North Carolina State University, Raleigh, NC 27695-7918

diamond turning, even for machining with a sharp tool [9]. The shear and normal forces on the shear zone inclined at the angle ϕ are F_s and F_n , respectively, while μ is the coefficient of friction between the tool and the chip. The equation for the total cutting force is obtained through a balance of forces in the horizontal direction. The total normal or thrust force, F_t , is provided by frictional forces on the tool rake face and the normal force associated with spring back ($F_{t\text{elastic}}$). Thus, the force components are given by

$$F_c = F_s \cos \phi + F_n \sin \phi + F_{c\text{elastic}} \quad (1)$$

$$F_t = \mu(F_c - F_{c\text{elastic}}) + F_{t\text{elastic}} \quad (2)$$

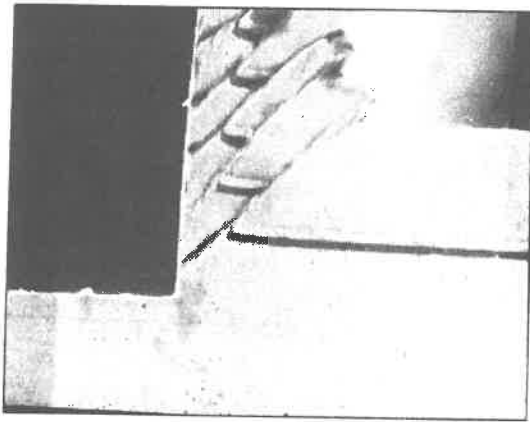


Figure 1. Cross section of plated copper chip.

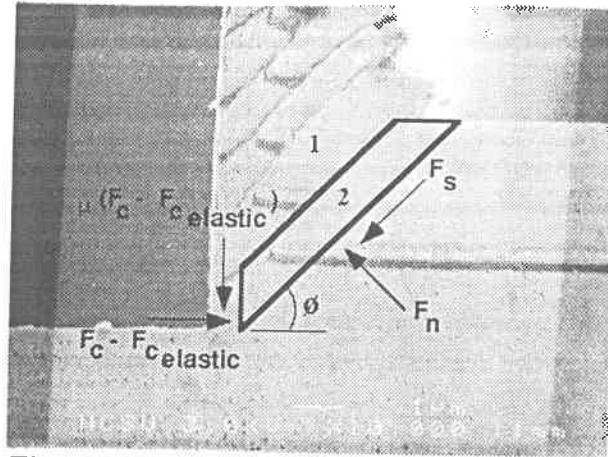


Figure 2. Chip cross section and free body diagram of chip segment.

Assumptions. Equations (1) and (2) contain several assumptions. The previously formed segment is regarded as exerting no shear force on the workpiece at the instant when a new segment forms. The material is regarded as elastic-perfectly plastic. The normal stress acting on the shear zone, σ_s , is taken as the uni-axial material flow stress, or one third the material hardness. Finally, a von Mises failure criterion is employed so that the shear stress, τ_s , is equal to $\sigma_s / \sqrt{3}$. These last two assumptions provide the maximum allowable normal and shear stresses in the cases of pure tension and pure shear, respectively. In a material that has a constant yield condition, and in which shear and normal stresses exist simultaneously, the magnitude of each component at yielding would be less than σ_s and τ_s as described in the preceding sentences.

Tool Force Equations. With these assumptions and the measured shear angle, ϕ , the coefficient of friction at the tool/chip interface, μ , can be calculated from a summation of forces in the vertical direction. The result of this operation yields

$$\mu = \frac{\cos \phi - \frac{\sin \phi}{\sqrt{3}}}{\frac{\cos \phi}{\sqrt{3}} + \sin \phi} \quad (3)$$

If the material hardness and either the shear angle or the coefficient of friction are known, the tool forces can be calculated. The development of models for material spring back and the associated contact stress between the tool flank and the workpiece is discussed by the authors elsewhere [9]. This work demonstrated that spring back in metal machining can be represented as a linear function of tool edge radius and the ratio of material hardness to elastic modulus. Thus, spring back is given by

$$s = k_1 r \frac{H}{E} \quad (4)$$

where s is the deflection of the machined surface or spring back, k_1 is a constant, r is the tool edge radius, H is the material hardness and E is the modulus of elasticity. For the 0° rake tool and the materials studied, $k_1 = 43$ provides good correlation between the measured and calculated values for spring back [9]. The contact stress due to this elastic spring back, σ_f , has been modeled as

$$\sigma_f = k_2 H \sqrt{\frac{H}{E}} \quad (5)$$

where k_2 is a constant ($k_2 = 4.1$ for best agreement with measured data [9]).

Knowledge of spring back allows the tool flank/workpiece contact area to be calculated. The product of this area and the interface stress yields the thrust force associated with elastic deformation of the workpiece. In terms of material data and cutting conditions, the complete tool force equations are then

$$F_c = \frac{H}{3\sqrt{3}} \frac{A_c}{\sin \phi} \cos \phi + \frac{H}{3} A_c + F_{c\text{elastic}} = \frac{HA_c}{3} \left(\frac{\cot \phi}{\sqrt{3}} + 1 \right) + \mu_f A_f \left(4.1 H \sqrt{\frac{H}{E}} \right) \quad (6)$$

$$F_t = \mu (F_c - F_{c\text{elastic}}) + F_{t\text{elastic}} = \mu \left[\frac{HA_c}{3} \left(\frac{\cot \phi}{\sqrt{3}} + 1 \right) \right] + A_f \left(4.1 H \sqrt{\frac{H}{E}} \right) \quad (7)$$

where A_c is the uncut chip area, A_f is the contact area between the tool flank and the workpiece and μ_f is the coefficient of friction between the tool flank and machined surface. Due to dissimilar conditions which exist along the nose region of the tool, the coefficient of friction between the tool flank and the workpiece is not required to be equal to the rake friction coefficient that is calculated from Equation (3) [10, 11].

EVALUATION OF THE MODEL

To verify the force model, the depth of cut, feed, tool nose radius, tool edge geometry, material properties (H , E) and either the shear angle or the coefficient of friction between the chip and tool rake must be known. Material data is provided in Table 1. The Vickers hardness of a diamond turned sample of each material was measured with a Zwick 3212 hardness tester. The elastic moduli were taken from a handbook [12]. In performing the calculations, the measured value for the coefficient of friction between the tool flank and the machined surface was used. For the four materials studied, this friction coefficient is about 0.2 [10]. Table 1 also contains shear angle data for the various materials and tool sharpnesses. The edge geometry of the diamond tool was varied by machining 6061-T6. The freshly lapped diamond tool had an edge radius of about 50 nm. The “worn tool” parameters (edge radius, flank wear land length) are also given in Table 1.

Material	Vickers Hardness (GPa)	Elastic Modulus (GPa)	Shear Angle (new tool)	Shear Angle (worn tool)
electroless nickel	5.3	200	22°	16° (300 nm, 2.4 μ m)
plated copper	2.4	117	45°	33° (300 nm, 2.4 μ m)
6061-T6	1.0	69	40°	30° (270 nm, 2.2 μ m)
5086	1.0	71	30°	22° (270 nm, 2.2 μ m)

Table 1. Material and machining data.

Comparisons of measured and calculated tool forces as a function of uncut chip area for electroless nickel cut with the 0° rake diamond tool in both its sharp and worn conditions are shown in Figures 3 - 6. The uncut chip area is the projection of the tool into the workpiece in the direction of cutting. The cutting speed for these tests was 25 mm/min. Figure 3 demonstrates excellent correlation between the measured and predicted forces in the cutting direction. The model accurately describes the tool forces for very large values of uncut chip area, A_c , as well as those typically encountered in diamond turning ($A_c < 400 \mu\text{m}^2$). The model also provides a good estimate of the thrust force over a large range of uncut chip area, as shown by Figure 4. For the worn tool geometry, good correlation between the measured and predicted cutting forces was also obtained, as demonstrated by Figure 5. However, the thrust forces (Figure 6) are overestimated by an average of 50%. The high values of calculated thrust force in this case are due to the imperfect model for flank stress associated with spring back. As the tool sharpness decreases, the forces due to elastic deformation form a larger percentage of the total force[12]. Thus errors in the model for the flank stress (Equation (5)) become more significant with tool wear.

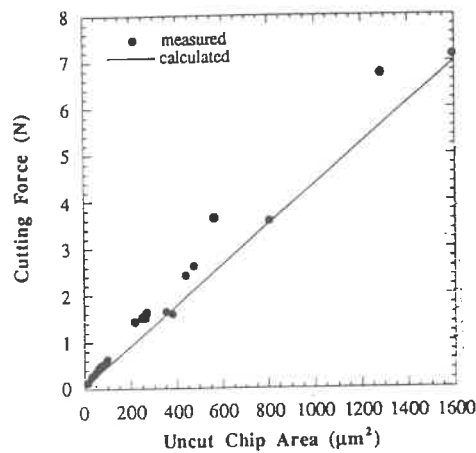


Figure 3. Measured and calculated cutting forces for electroless nickel (sharp tool).

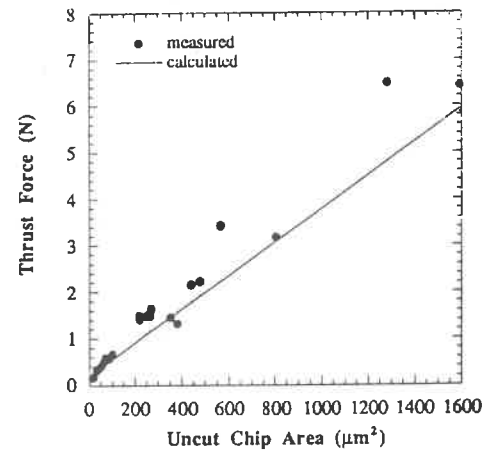


Figure 4. Measured and calculated thrust forces electroless nickel (sharp tool).

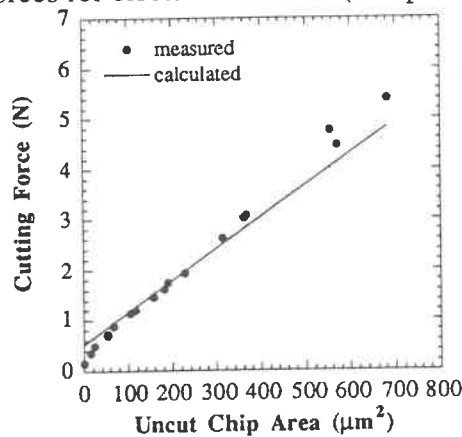


Figure 5. Cutting forces for electroless nickel (300 nm edge radius, 2.4 μm land).

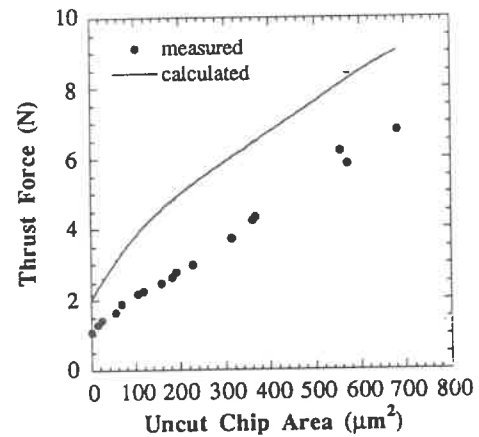


Figure 6. Thrust forces for electroless nickel (300 nm edge radius, 2.4 μm land).

Using the data in Table 1, the force model was evaluated for the three remaining materials and two tool conditions. The average variations between the measured and calculated forces in the cutting direction for chip areas less than 800 μm^2 are indicated in Table 2 (see Table 1 for worn tool data). Good correlation was also obtained for the thrust forces.

Material	Avg. Deviation from Measured Force (sharp tool)	Avg. Deviation from Measured Force (worn tool)
Plated Copper	25% (underpredicted)	30% (underpredicted)
6061-T6	30% (underpredicted)	20% (underpredicted)
5086	<10%	<10%

Table 2. Comparison of cutting force data.

CONCLUSION

A tool force model that is applicable in precision machining has been developed. This model accounts for chip formation as well as the elastic deformation of the workpiece that does not lead to material removal. The force components in the cutting and thrust (normal) directions are calculated

from the material hardness, elastic modulus, uncut chip area, tool geometry and shear angle for chip formation. The model was verified through the low speed machining of electroless nickel, electroplated copper and two aluminum alloys with a 0° rake angle diamond tool. These cutting tests were performed over a wide range in chip area and tool sharpness. The correlation between the measured and predicted force components is excellent for machining with a sharp tool. This agreement, however, generally degrades for cutting with a worn tool. The discrepancy between the measured and calculated forces for machining with a dull tool is primarily due to the increasing significance of elastic deformation of the workpiece with tool wear and inaccuracies in modeling the tool force caused by this elastic deflection. Despite this limitation, the predicted thrust forces vary by less than 50% from those measured for all the materials. The cutting forces are in better agreement, differing from the measured values by less than 40% for all materials and tool conditions. The relationships between cutting parameters and machining force that are incorporated in this model should help to establish tool force as a valuable process monitor in precision machining.

REFERENCES

1. Finnie, I. "Review of Metal Cutting Analyses of the Past Hundred Years." *Mechanical Engineering*, v. 78, 1956: 715 - 721.
2. Ernst, H. and Merchant, M. E. "Chip Formation, Friction, and High Quality Machined Surfaces." *Surface Treatment of Metals*, ASM, 1941: 299 - 335.
3. Shaw, M.C.; Cook N.H. and Finnie, I. "The Shear-Angle Relationship in Metal Cutting." *Transactions of the ASME*, v. 75, 1953: 273 - 288.
4. Lee, E.H. and Shaffer, B.W. "The Theory of Plasticity Applied to a Problem of Machining." *J. of Applied Mechanics*, v. 18, 1951: 405 - 413.
5. Rowe, G.W. and Spick, P.T. "A New Approach to Determination of the Shear-Plane Angle in Machining." *J. of Engineering for Industry*, v. 89, 1967: 530 - 538.
6. Arcona, C. and Dow, T.A. "A New Technique for Studying the Chip Formation Process in Diamond Turning." *Precision Engineering*, v. 18, n. 2/3, 1996: 157 - 160.
7. Black, J.T. "Flow Stress Model in Metal Cutting." *J. of Engineering for Industry*, v. 101, 1979: 403- 415.
8. Ueda, K. and Manabe, K. "Chip Formation Mechanism in Microcutting of an Amorphous Metal." *Annals of the CIRP*, v. 41, part 1, 1992: 129 - 132.
9. Arcona, C. and Dow, T.A. "Elastic Components of Deformation and Tool Force in Precision Machining." *Precision Engineering*, in press.
10. Arcona, C. "Tool Force, Chip Formation and Surface Finish in Diamond Turning." Ph.D. Dissertation. North Carolina State University, 1996.
11. Enahoro, H.E. and Oxley, P.L.B. "Flow Along Tool-Chip Interface in Orthogonal Metal Cutting." *J. of Mechanical Engineering Science*, v. 8, n. 1, 1966: 36 - 41.
12. Materials Engineering, *Materials Selector 1989*. Penton Publishing, Cleveland, OH, December, 1988.

Effect of Grain Orientation and Grain Boundaries on the Surface Topography of Diamond Turned OFHC Copper

E. Brinksmeier, W. Preuss, O. Riemer, J. Schmütz
LFM Labor für Mikrozerspanung, University of Bremen, Germany

1 Introduction

The microtopography of diamond turned polycrystalline metals depends on cutting parameters and on the crystallographic orientation of individual grains [1-4]. We are studying the influence of grain orientation on the microroughness of OFHC copper grains obtained in diamond cutting experiments. The aim of our work is to classify microtopographic features and to establish a relationship between grain orientation, cutting forces and surface topography.

2 Experimental setup

Inclined plunge-cut experiments with increasing depth of cut were performed in coarse-grained OFHC copper. The cutting experiments were carried out on a Moore M18 Aspheric Generator [2]. The diamond tools were mounted on the clamped spindle fixture. The copper specimen were slightly tilted (approx. 1 : 1000) and fixed on a piezo-electric force transducer. The relative motion between tool and specimen was obtained by moving the x-axis table at 10 mm/min and 100 mm/min, respectively.

Figure 1 shows the set-up of the inclined plunge-cut experiments.

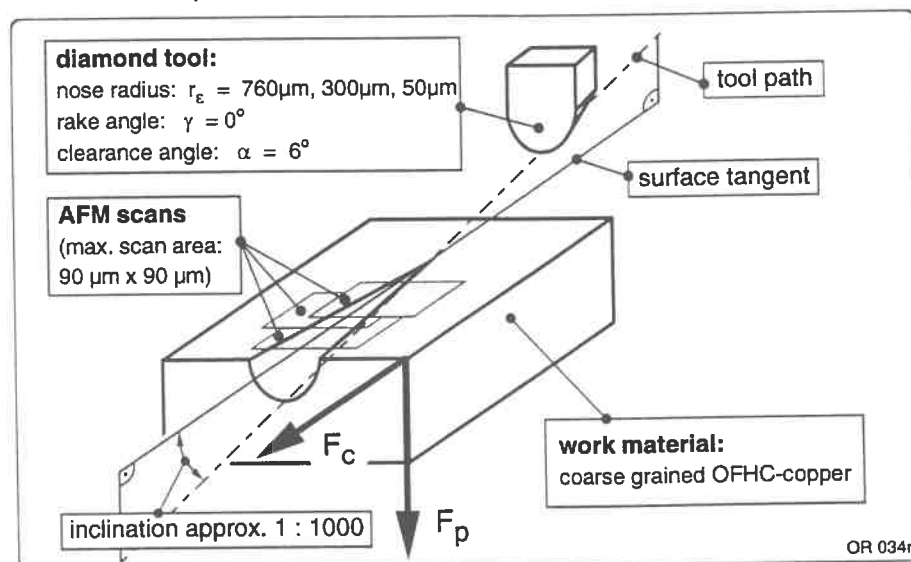


Fig. 1: Set-up of inclined plunge-cut experiments.

Each plunge-cut of the test series was repeated at least five times. The plunge-cut grooves were fully scanned and documented by Normarski microscope pictures.

Topographic details were studied with an atomic force microscope (AFM). Three different diamond tools with a nose radius of 760 μm , 300 μm to 50 μm were used. All tools had a 0° rake angle and a 6° clearance angle. The cutting edges radii of the tools, examined with a Nanoscope III atomic force microscope, were determined to be less than 30 nm.

3 Recrystallization of OFHC copper

Fine-grained OFHC copper commonly used for micromachining has an average grain size of approximately 80 μm . Single grains of this fine-grained polycrystalline material cannot be investigated by X-ray diffraction for determining crystal orientation, since the focus diameter is 0.1 mm or larger. Therefore, recrystallization was induced by cold rolling and annealing.

Since full recrystallization diagrams of copper do not exist, different samples made from commercially available OFHC copper were cold rolled with deformation ratios varying from 5 % to 15 %. Finally, a specimen was chosen which was cold rolled with a deformation ratio of 10 % followed by tempering at 830 $^\circ\text{C}$ for one hour in a nitrogen atmosphere. This process yielded a homogeneous grain size distribution with an average size of 330 μm . Though recrystallization should have no influence on the mechanical properties of the sample, the grain size does affect hardness testing. Therefore, the micro-hardness of the fine-grained samples (123 HV 0,05) was greater than the hardness after rolling and annealing (87 HV 0,05).

4 Experimental results

Analyzing the surface texture generated by plunge-cutting, different microtopographies were found depending on the depth of cut a_p . As long as $a_p < 2 \mu\text{m}$, all grooves exhibit a very smooth micro-topography (cf. picture on the left in **figure 3**) whatever the crystal orientation, and no extraordinary features were found at grain boundaries. A shallow waviness is observed along the cutting direction probably caused by the chip formation process. The average roughness R_a of these areas is less than 10 nm. The same observation is made with different diamond tools and cutting speeds. Moreover, plunge-cuts into fine-grained copper with $a_p < 2 \mu\text{m}$ also show the same smooth surfaces whose topography seems to be characteristic for undisturbed machining.

If the depth of cut $a_p > 2 \mu\text{m}$, sections characterized by increased roughness and a markedly different microtopography appear in all grooves. These areas generally extend over a complete grain, but in some cases start out at a grain boundary and disappear before reaching the next boundary. **Figure 2** shows three adjacent plunge-cut grooves cut under identical experimental conditions, exhibiting both smooth areas and areas with increased roughness.

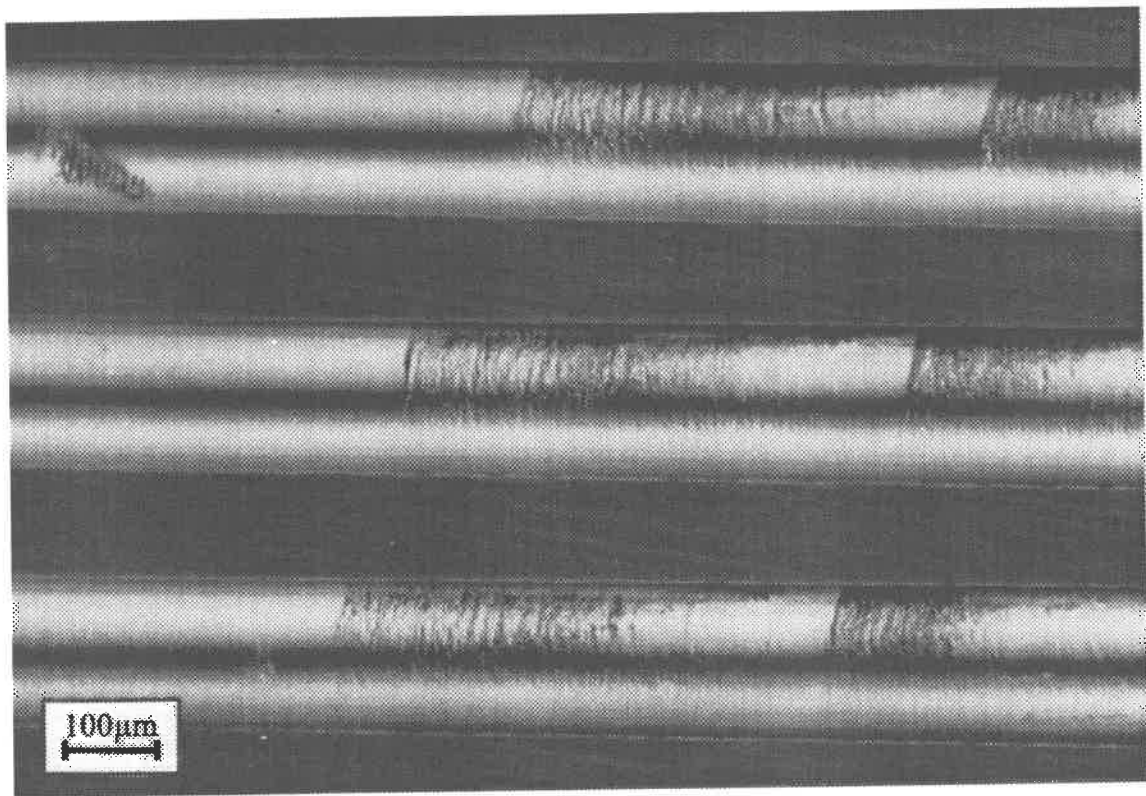


Fig. 2: Normarski photograph of three adjacent plunge-cut grooves. Tool nose radius: $r_e = 300 \mu\text{m}$, cutting speed: $v_c = 10 \text{ mm/min}$, cutting direction: left to right.

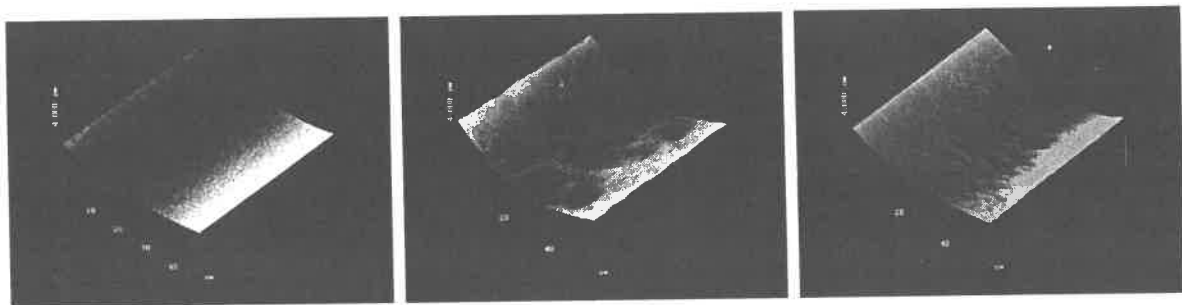


Fig. 3: Smooth, dented and rippled sections at the bottom of plunge-cut grooves. Scale: $50 \mu\text{m} \times 50 \mu\text{m}$, lateral; $4 \mu\text{m}$, vertical. Tool nose radius: $r_e = 300 \mu\text{m}$, cutting speed: $v_c = 10 \text{ mm/min}$, cutting direction: left to right.

We distinguish two types of areas with increased roughness, i. e. those with dented and others with rippled features. Examples of the different types of microtopography are shown in **figure 3** which can be characterized as follows:

- smooth sections: - $R_a < 10 \text{ nm}$
- always generated if depth of cut $a_p < 2 \text{ }\mu\text{m}$
- dented sections: - $10 \text{ nm} < R_a < 100 \text{ nm}$
- exhibiting small dents of approx. $10 \text{ }\mu\text{m}$ in diameter
- rippled sections: - $R_a > 100 \text{ nm}$
- exhibiting ripples perpendicular to the cutting direction

All possible transitions between these three different types of surface topography can be found in the plunge-cut grooves. Thus, microtopography is correlated with individual grain orientation and is not exclusively depending on depth of cut.

If dented and rippled sections are regarded as a surface degradation, the ratio of degraded to total surface area is increasing rapidly with increasing depth of cut. Degradation up to 90 % can be found at larger depths of cut. Also, cutting force signals split into two or three different levels which can be correlated with microtopography. Areas with increased roughness show a considerably higher cutting force level than smooth sections. At grain boundaries abrupt changes between the force levels occur.

5 Conclusion

Inclined plunge-cut experiments with increasing depth of cut were performed in coarse-grained OFHC copper. Microtopographic features were divided into three distinct classes and correlated with cutting force measurements. It is assumed that the observed surface patterns depend on the number of gliding systems activated in the microcutting process. The analysis of X-ray diffraction patterns obtained from individual grains will provide a quantitative proof of this assumption.

6 References

- [1] Brinksmeier, E. "Herstellung von Präzisionsbauteilen durch Mikrozerspanung". Industrie Diamanten Rundschau 4/1994: S. 210-217.
- [2] Brinksmeier, E; W. Preuß; O. Riemer. „From Friction to Chip Removal. An Experimental Investigation of the Microcutting Process. Part I“. Proc. 3rd Int. Conf. on Ultra Precision in Manufacturing Engineering; Hrsg: M. Weck, H. Kunzmann, Verlag Franz Rhiem, Duisburg, 1994: p. 393-400.
- [3] Rentsch, R. „Molecular dynamics simulation for nanometer chip removal processes“. Ph. D. thesis, Keio University, Japan, 1995.
- [4] Spenrath, N. „Technologische Aspekte zum Feinstdrehen von Kupferspiegeln“. Dissertation, RWTH-Aachen, 1991.

Design and Application of In-Process Acoustic Sensors for Maximum Sensitivity-Methods and an Example in Diamond Tooling

Jay Daniel, David Dornfeld, Laboratory for Manufacturing Automation, Department of Mechanical Engineering, University of California, Berkeley

Abstract: Acoustic sensors have been shown to be of value in precision manufacturing as near contact and contact sensors in a number of machining applications. Acoustic sensors have also shown promise in the detection and discrimination of various cutting parameters in cutting processes. Design methods that allow the sensitivity of these sensors to be maximized are not commonly available though, forcing users to adapt commercial acoustic emission equipment optimized to no particular situation. In particular, manufacturers are required to adapt existing sensors to the geometric constraints of their machining setup, and doing so can be difficult in many situations, restricting the utility of this sensing methodology. A quantitative design methodology that allows the manufacturer to consider tradeoffs in sensitivity and geometry for a unbacked bulk wave transducers is given. A design example for an acoustic sensor for diamond turning is considered.

INTRODUCTION

Acoustic sensors have been frequently investigated as in-process sensors for precision manufacturing processes.^{1,2,3} They have been shown to be particularly robust in detecting changes in cutting condition.⁴ Unfortunately, investigators seeking to incorporate acoustic emission sensors into process control schemes normally have had to work with the limits of commercially available acoustic emission transducers. These transducers can be limiting in terms of their size and cost.

In addition, maximizing sensitivity of acoustic receivers has not always driven design decisions for sensor manufacturers. With the appropriate models at hand, the ability to "design in" a transducer into a manufacturing process allows the manufacturing engineer to maximize the utility of an acoustic sensor. This is particularly true in precision manufacturing applications, where exceedingly high sensitivity and signal to noise ratio is required for all process signals.

Previous authors have developed parameters ranking the sensitivity of acoustic receivers.^{5,6} These models were not derived from first principles, and also did not incorporate the effects of the substrate through which the acoustic emission was propagating, or the parameters of the signal amplification. All of these parameters should be considered in designing an acoustic receiver for a particular machining application. The first step in considering piezoelectric acoustic receivers is the development of an appropriate model.

Following this, critical aspects of the model can be applied to given engineering situations to maximize sensor sensitivity. An example of this type of consideration is given for selecting a piezoelectric acoustic receiver for diamond tooling.

BULK WAVE MODEL FOR PIEZOELECTRIC RECEIVERS

Our goal is to develop an expression for the voltage generated by a piezoelectric receiver as it is acted upon by an acoustic wave propagating in from a substrate that it is mounted upon. Figure 1 shows the physical situation, and some of the terms used in the model.

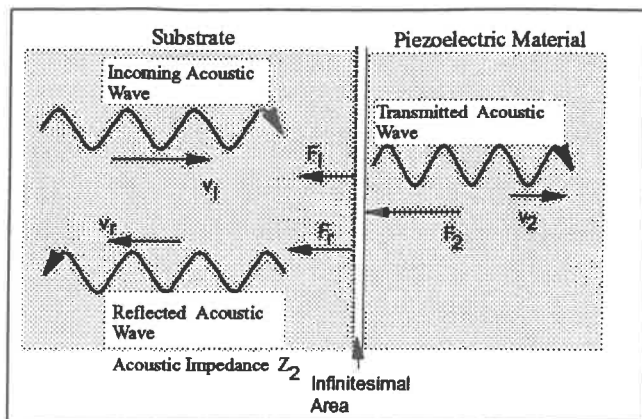


Figure 1: The physical model for an acoustic receiver showing the acoustic waves and traction and velocity vectors.

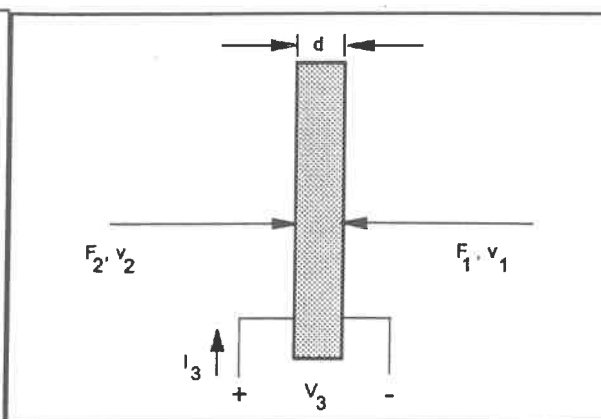


Figure 2: The three port model and terms used in the Mason model.

The transfer function we wish to obtain is the voltage generated by the piezoelectric material as generated by some aspect of the incoming acoustic wave. For our purposes, we choose particle displacement. Thus we want the transfer function V_3/u_i where V_3 is the voltage generated across the piezoelectric material.

To obtain this transfer function, we utilize the Mason model⁷ for plane wave propagation in piezoelectric media along with relationships for plane wave reflection at an interface. The Mason model is

$$\begin{bmatrix} F_1 \\ F_2 \\ V_3 \end{bmatrix} = -i \begin{bmatrix} Z_o \cot(kd) & Z_o \csc(kd) & \frac{h}{\omega} \\ Z_o \csc(kd) & Z_o \cot(kd) & \frac{h}{\omega} \\ \frac{h}{\omega} & \frac{h}{\omega} & \frac{1}{\omega C_o} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ I_3 \end{bmatrix} \quad (1)$$

and relates the forces and velocities at the front and back faces of the piezoelectric plate to physical parameters of the piezoelectric media and the voltage and current generated across the piezoelectric plate. Terms used in the model include the dynamic acoustic impedance(Z_o), the piezoelectric constant(h), the wave vector(k), the plate thickness(d), and the plate capacitance(C_o). Figure 2 shows the geometry of the model. For the excitation of an unbacked transducer from the substrate, we apply two boundary conditions to the Mason Model.

$$F_1 = 0 \quad (2)$$

and

$$V_3 = -Z_3 I_3 \quad (3)$$

The first boundary condition is used because the transducer is unbacked, and no force can be supported at the back of the plate. The second boundary condition applies to the electrical terminal. We assume that an impedance is placed across the electrical terminal across which the voltage generated by the piezoelectric plate is measured.

Finally, an expression needs to be developed that describes how much of an acoustic wave impinging on the piezoelectric plate from the substrate propagates into the piezoelectric plate, and how much reflects back into the substrate. First, we state the reflection coefficient of an interface, Γ , as a function of the impedances.⁸ Z_{in} is the acoustic impedance looking into the piezoelectric plate from the substrate.

$$\Gamma = \frac{Z_{in} - Z_2}{Z_{in} + Z_2} \quad (4)$$

Also, the traction vector at the boundary of the piezoelectric plate, F_2 , is related to F_i and F_r by:

$$F_2 = F_i + F_r = F_i (1 + \Gamma) = F_i \frac{2Z_{in}}{Z_{in} + Z_2} \quad (5)$$

Three other relationships need to be used to find V_3/u_i .

$$F_i = -\bar{Z}_2 v_i \quad (6)$$

$$Z_{in} = \frac{F_2}{v_2} \quad (7)$$

$$u_i = \int v_i dt = \int v_{io} e^{-i(kx - \omega t)} dt = \frac{v_{io}}{i\omega} e^{-i(kx - \omega t)} = \frac{v_i}{i\omega} \quad (8)$$

Using this system of equations, we can now solve for V_3/u_i .

$$\frac{V_3}{u_i} = \left[\frac{Z_3 Z_2}{\frac{Z_o}{h} (\cot(kd) + \csc(kd)) \left(iZ_3 + \frac{1}{\omega C_o} \right) - 2 \left(\frac{h}{\omega} \right)^2} \right] \begin{bmatrix} \frac{2 \left(\frac{h}{\omega} \right)^2}{\left(iZ_3 + \frac{1}{\omega C_o} \right) Z_o} (\sec(kd) - 1) - \tan(kd) \\ \frac{\left(\frac{h}{\omega} \right)^2}{\left(iZ_3 + \frac{1}{\omega C_o} \right) Z_o} \tan(kd) - 1 \end{bmatrix} Z_o i \quad (9)$$

$$\begin{bmatrix} \frac{2 \left(\frac{h}{\omega} \right)^2}{\left(iZ_3 + \frac{1}{\omega C_o} \right) Z_o} (\sec(kd) - 1) - \tan(kd) \\ \frac{\left(\frac{h}{\omega} \right)^2}{\left(iZ_3 + \frac{1}{\omega C_o} \right) Z_o} \tan(kd) - 1 \end{bmatrix} Z_o i + Z_2$$

Using this model, the effects of design parameters including the thickness and area of the piezoelectric material, the piezoelectric material and its properties, and the input impedance of the amplifier can be evaluated.

Assuming that Z_3 , the input impedance of the amplifier used to amplify the signal from the piezoelectric receiver, is infinite, (9) reduces to⁹

$$\frac{V_3}{u_i} = \frac{-2hi}{\frac{Z_o}{Z_2} \{ \cot(kd) + \csc(kd) \} \left\{ 1 + \frac{Z_2 i}{Z_o} \cot(kd) \right\}} \quad (10)$$

The figure of merit for various materials can be deduced from (10). The figure of merit (FOM) is the same well below and near resonance of the transducer, although the transducer acts as a surface velocity transducer well below resonance and a surface displacement transducer near resonance.⁹ The figure of merit is

$$FOM = \frac{hZ_2}{Z_o} \quad (11)$$

Table 1 lists a variety of piezoelectric materials by their figure of merit. The substrate impedance, Z_2 , isn't included in the FOM, as it isn't a transducer property. It does have tremendous importance in the sensitivity of an acoustic detection system, but often isn't controllable. Table 1 lists the relative permittivity of the piezoelectric material, the piezoelectric stress constant, the specific dynamic impedance, and the FOM. The piezoelectric parameter, h , is related to the relative permittivity, ϵ , and the piezoelectric stress constant, e , by

$$h = \frac{e}{\epsilon} \quad (12)$$

Table 1 also includes a listing of whether the receivers are commonly used as thin film or macroscopic transducer materials.

A DESIGN EXAMPLE- DIAMOND TOOLING FOR A LATHE

Using these design guidelines, two sensor configurations were evaluated for an acoustic sensor mounted on a diamond tool used in a turning center. One configuration utilizes a macroscopic piece of PZT with a commercially available amplifier, while another utilizes a thin film zinc oxide sensor with a custom high frequency amplifier. Figure 3 demonstrates the locations of the sensors on the diamond tool.

MATERIAL	relative permittivity no units	piezo-stress const C/m2	Spec. Dynamic Impedance kg/(sec.m ²)	Figure of merit kg.C/(F.sec.m ⁴ 5)	Thin Film Material? Y/N/both
SiO2	4.5	0.50	1.33E+07	982	N
AlN	8.5	2.12	3.40E+07	858	Y
Kynar	12	0.32	3.90E+06	792	Y
CdS	9.5	1.69	3.60E+07	580	Y
LiNbO3	39	6.56	3.42E+07	576	N
Keramos K83	150	8.90	2.56E+07	271	N
ZnO	8.8	0.70	3.60E+07	260	both
Murata SVM	240	13.44	3.74E+07	175	N
PZT-7A	235	9.5	3.65E+07	130	N*
PZT-2	260	9	3.35E+07	121	N
Na0.5K0.5NbO	306	9.8	3.09E+07	121	N
P3 BaTiO3	890	24.50	3.13E+07	103	both
P6 PZT	890	25.00	3.51E+07	94	N
P7 PZT	1000	27.57	3.60E+07	90	N
P5 PZT	850	19.25	3.16E+07	84	N
PZT-4	635	15.1	3.45E+07	81	N
PZT-8	580	13.2	3.48E+07	77	N
PZT-5A	830	15.8	3.37E+07	66	N
PZT-6B	386	7.1	3.64E+07	59	N
PZT-6A	730	12.5	3.40E+07	59	N
BaTiO3	1260	17.5	3.12E+07	52	both

Table 1: Ranking of piezoelectric materials as receivers. Data from references 8 and 10.

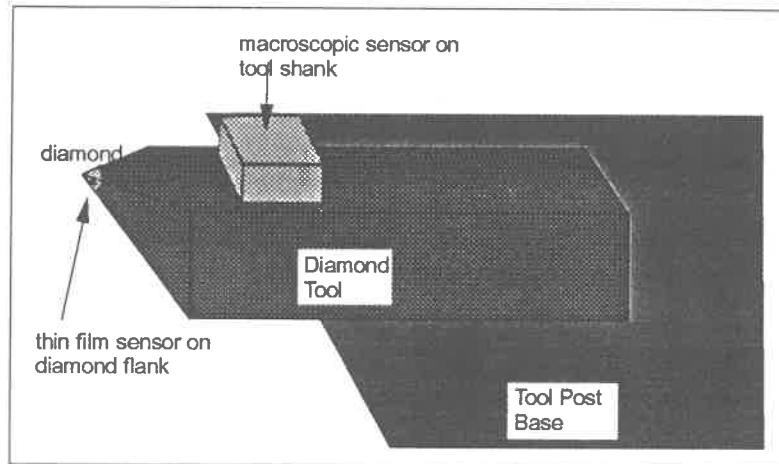


Figure 3: Locations of acoustic sensors on diamond tooling.

The considerations in maximizing sensitivity in this design example are the relative sensitivities of the acoustic sensors and the intensity of the acoustic waves at the two sensor locations. The sensitivities of the two sensor configurations can be calculated with (9) over the frequency range monitored. The ZnO sensor is significantly less sensitive than the PZT sensor, due to its operation well below the resonances of the sensor. The magnitude of the displacements generated by acoustic waves propagating as a result of time varying tractions imposed on the tool tip by the cutting process can be nominally estimated by FEM,⁹ but the frequency capability of the FEM analysis is limited by element size in the FEM mesh. Results from the FEM analyses show that the macroscopic transducer configuration should be more sensitive overall, but the level of confidence in this result is not high due to the limits of the FEM analysis. The testing of these two sensor configurations is currently ongoing in our laboratory, and results will be published in the next year.

CONCLUSIONS

A model for the sensitivity of plane wave piezoelectric acoustic transducers has been developed. The figure of merit for various piezoelectric acoustic receivers has been extracted from the model. Estimates of sensitivity of two acoustic sensor configurations for an application in diamond tooling have been obtained, and experiments are being conducted to compare experimental results to the theoretical predictions.

- ¹ Taylor, J.S., et. al., Proceedings of the ASPE, Austin, TX, USA, Nov. 1995
- ² Osuri, R.H., et. al., Int. J. of Prod. Research, 29(7), 1991.
- ³ Kamimura, Y., Tani, Y., Proc. 8th Int. Precision Eng. Sem., Compiegne, France May 1995.
- ⁴ Dornfeld, D.A., JSME Int. Journal C, 34(4), 1994.
- ⁵ Callerame, J., Tancrell, R.H., Wilson, D.T., 1979 IEEE Ultrasonics Symposium.
- ⁶ Silk, M.G., *Ultrasonic Transducers for Non-Destructive Testing*, Adam Hilger Ltd., Bristol, 1984.
- ⁷ Mason, W.P., *Piezoelectric Crystals and Their Application to Ultrasonics*, van Nostrand, New York, 1950.
- ⁸ Auld, B., *Acoustic Fields and Waves in Solids*, 2nd ed., Robert E. Krieger Publishing Co., Malabar, FL.
- ⁹ Daniel, J.P. Ph.D. Thesis, University of California, Berkeley, 1996.
- ¹⁰ Kino, G.S., *Acoustic Waves. Devices, Imaging, & Analog Signal Processing*, Prentice Hall, Englewood Cliffs, New Jersey, 1987.

Development of a Micro-Cutting/Testing System under Scanning Electron Microscope

Mamoru MITSUISHI, Naohiko SUGITA, Takaaki NAGAO and Yotaro HATAMURA

Department of Engineering Synthesis, Faculty of Engineering

The University of Tokyo

Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan

mamoru@nml.t.u-tokyo.ac.jp

<http://www.nml.t.u-tokyo.ac.jp/~mamoru>

1. Introduction

Etching processes and the LIGA process are used conventionally to fabricate 3D micro-mechanical parts. It is also possible to manufacture 3D mechanical parts by cutting, using micro-tools. In this paper, a micro-machining system which works under an SEM was developed. The system incorporates 3-axis force sensing to measure and analyze cutting force. From the experiment it determined that as the cutting scale becomes smaller, the cutting force per unit increases. Therefore, a high stiffness micro tool is necessary for micro cutting.

2. Related work

An example of research on micro-machining and testing is micro-operation within an SEM [1]. The system includes a specimen stage and a 3 degree-of-freedom tool-handling robot, equipped with a single axis force sensor. It is driven by piezoelectric actuators and is installed in the vacuum chamber of the SEM. Johansson [2] has realized a system which can test the physical

properties of a material, such as diffusion, at 500 degrees centigrade by holding a specimen in the SEM chamber. Takeuchi [3] has developed a lathe-type micro-milling machine. Brittle material such as glass was machined in the ductile mode. It was possible to obtain a glass mask of 1 mm in diameter with the surface roughness of 50 nm. Friedrich [4] has developed a micro-milling machine with milling tools with a diameter as small as 22 micrometers. The milling tools were made using FIB. Shimada[5] and Shimizu[6] are analyzing the micro cutting mechanism using the knowledge of molecular dynamics. Furthermore, Tokushige[7] is executing research concerning micro cutting in the vacuum environment.

3. Construction of a machine for micro-machining within an SEM

3.1 Construction of a micro-machining machine

Fig.1 shows the construction of a micro-machining machine (MMM) with three degrees

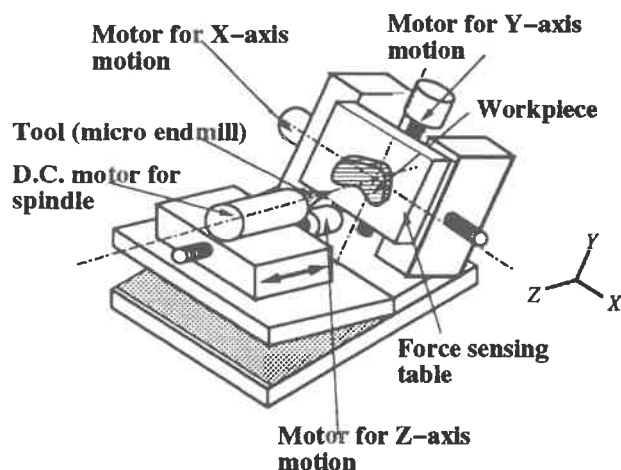


Fig.1 Construction of a micro-machining machine (MMM)

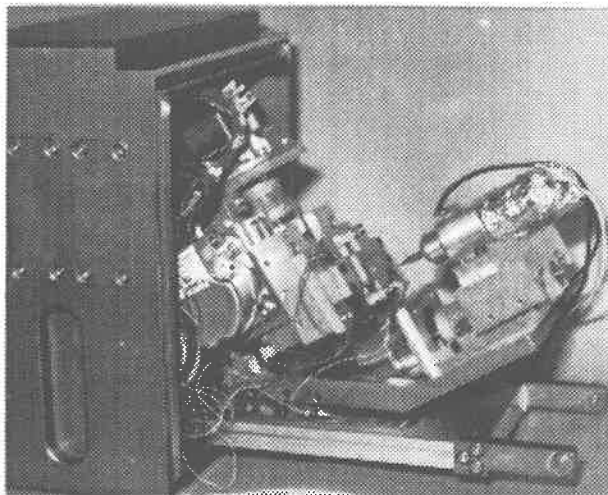


Fig.2 Overview of a micro-machining machine(MMM)

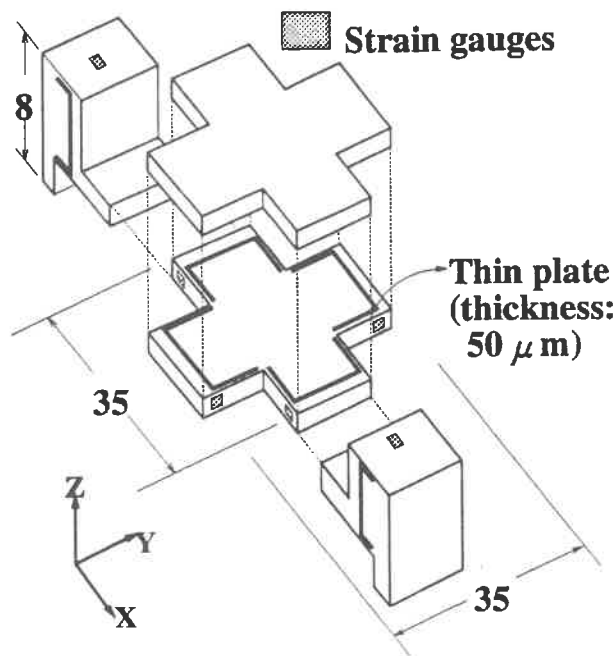


Fig.3 Structure of three-axis force sensing table

Table 1 Characteristics of the force sensing table

	Rated max. load value	Resolution	Natural frequency
X-axis	520mN	0.130mN	142Hz
Y-axis	752mN	0.188mN	104Hz
Z-axis	3528mN	0.882mN	324Hz

of freedom. The workpiece table translates in the X and Y directions and the tool moves in the Z direction. The translational mechanism uses ball screws and cross roller guides and is driven by ultrasonic motors with encoders. The working table is inclined at an angle of 60 degrees for optimal viewing under the SEM (Fig.2). The MMM has a rough positioning table to set the machining point at the center of the field of observation, with a three-axis force sensing table mounted on the XY-table (Fig.3). The sensing table is made of aluminum alloy and incorporates a parallel plate structure with strain gauge sensors. The thickness of the force sensing plate, which was produced using wire electrical discharge machining, is 50 μ m (length 4mm, width 3mm). Table 1 shows the characteristics of the developed force sensing table. For the MMM described in this paper, it is possible to fit a spindle motor on the Z-axis to rotate a small diameter end mill, which, when combined with Z-axis translation, produces a micro machining center. In our system, the shank portion of the

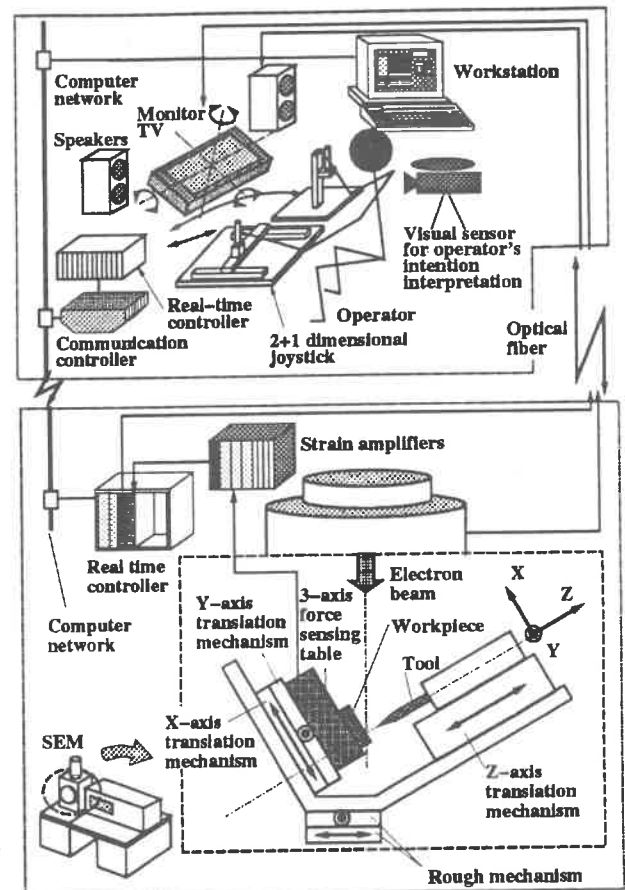


Fig.4 System configuration of a tele-micro machining system

endmill is directly supported by a bearing to reduce the deviation and the workpiece is fixed on the 3-axis force sensing table.

3.2 Tele-micro-machining system

The MMM is installed in the vacuum chamber of an SEM and micro machining is executed by controlling it with real-time controller which is connected to a computer network, such as the Internet (Fig.4). Visual information from the SEM and auditory information obtained the transformation of multi-axis force information into auditory information transformation method are sent through optical fiber to a remote operator. The remote operator controls the MMM, recognizing phenomena in the micro world by integrating the visual, auditory and tactile information. The control signal and force information are sent through a computer network, such as the Internet. For example, the MMM located at the University of Tokyo was operated from Karlsruhe University in Germany using ISDN as shown in Fig.5.

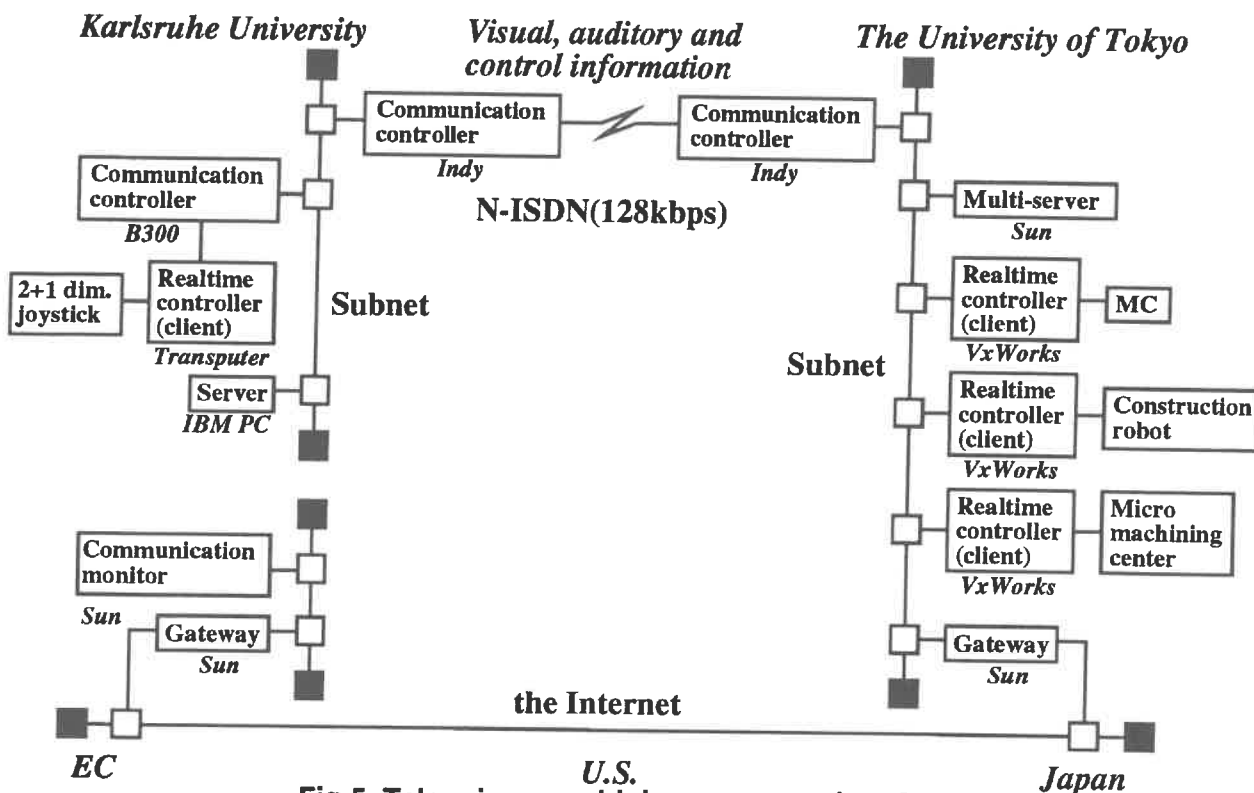


Fig.5 Tele-micro-machining system using ISDN

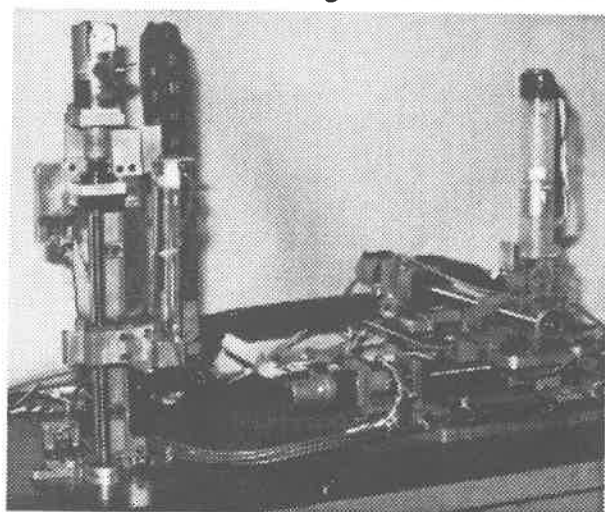


Fig.6 Overview of a 2+1 dimensional joystick

3.3 Auditory information presentation of machining state

Because the MMM is located inside the chamber of an SEM, it is difficult to acquire the cutting sound which is, in general, effective in indicating the machining state to a human operator. Therefore, the cutting state is detected by a multi-axis force sensor and the sensed information is converted into auditory information.

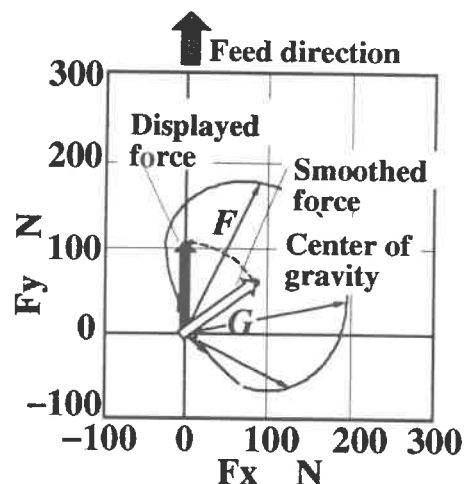


Fig.7 Cutting state index

Generally, the output from a force sensor is presented as follows for each force direction:

$$f(t) = \sum_{i=1}^{\infty} a_i \sin(\omega_i t + \theta_i) \quad (1)$$

where a_i , θ_i and ω_i are amplitude, phase and angular velocity, respectively. In the developed system the frequency of the sound ω_i was controlled to be proportional to the absolute value

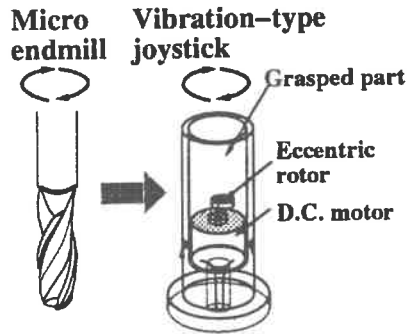


Fig.8 Vibration-type joystick

of the cutting force $\|f(t)\|$ while the amplitude A_i is constant.

$$\Phi(t) = A_i \sin((\alpha\|f(t)\| + \beta)t + \theta_i) \quad (2)$$

where α is a positive constant and β is defined as $2\pi \times 20$. The value 20 comes from the audible lower frequency limit of human hearing.

3.4 Tactile information presentation of machining state

In the implemented system, the machining state such as chatter vibration was presented as a vibration. To present the machining state the following value, named the "cutting state index" by the authors, is calculated in real-time and is sent from the micro-machining machine to the joystick controller:

$$\frac{V(|\mathbf{F} - \mathbf{G}|)}{|\mathbf{G}|^2} \quad (3)$$

where V , $\mathbf{F}$ and $\mathbf{G}$ represent the variance during one tool rotation, cutting force vector and the center of the gravity of cutting force during one tool rotation, respectively (Fig.7). The value $V(|\mathbf{F} - \mathbf{G}|)$ can be used to judge the machining state in real-time and the value becomes large when chatter vibration occurs [8]. However, the value changes depending on machining conditions such as feed per edge S_z even if the machining state is normal. Therefore, it is divided by $|\mathbf{G}|^2$ for the normalization. Fig.8 shows the grasp part of "2+1 dimensional joystick" with an eccentric weight and a motor at the operation room. The rotational speed of the motor is controlled as proportional to the value described in equation (3). The operator feels the vibration in case of chatter vibration.

4. Micro cutting experiments and discussions

Fig.9 shows one of the experimental results in micro machining. Fig.10 shows the plotted data

Table 2 Comparison of cutting force at macro-and-micro machining

	Micro cutting	Macro cutting	Magnification ratio
Tool diameter	0.2mm	10mm	$\alpha = 50$
Depth of cut	12.5 μm	5.0mm	$\beta = 400$
Feed per tooth	0.75 μm	25 μm	$\gamma = 33.3$
Maximum cutting resistance	100mN	200N	$\delta_{exp.} = 2 \times 10^3$

in the plane perpendicular to the main axis of a micro endmill, as obtained from the experimental data in Fig.9. It is clear that in micro cutting, in some cases the cutting state is normal and in other cases chatter occurred.

Cutting force per unit was compared in both micro and macro cutting cases under the conditions shown in Table 2. According to the Kronenberg equation, cutting force F can be written as follows:

$$F = k_c A \quad (4)$$

$$k_c = Ch^{-\epsilon} \quad (5)$$

$$h \simeq S_z \sin \phi \quad (6)$$

where k_c , A , h and S_z represent cutting force per unit area of the cut, area of cut, undeformed chip thickness and feed per edge, respectively. C and ϵ are constants which depend on the workpiece material and on the shape of cutter. Based on the value $\epsilon = 0.73$ obtained in macro cutting, the comparison ratio $\delta_{calc.}$ between macro and micro cutting can be calculated as follows, using variables shown in Table 2:

$$\delta_{calc.} = \alpha\beta\gamma^{-\epsilon} = 1.55 \times 10^3. \quad (7)$$

The value is the same order as the experimentally obtained comparison rate

$$\delta_{exp.} = 2.0 \times 10^3. \quad (8)$$

Fig.11 shows the plotted data describing the relationship between cutting force per unit area of the cut and undeformed chip thickness. The inclination of the straight line in the figure shows $-\epsilon$ which appears in an equation to calculate the cutting force. From Fig.11, it is easy to see that a stronger micro machine tool is necessary for micro cutting because the cutting force per unit increases as the scale becomes smaller. The experimental value obtained in micro cutting is approximately the same as the theoretical value 1.1×10^4 (MPa).

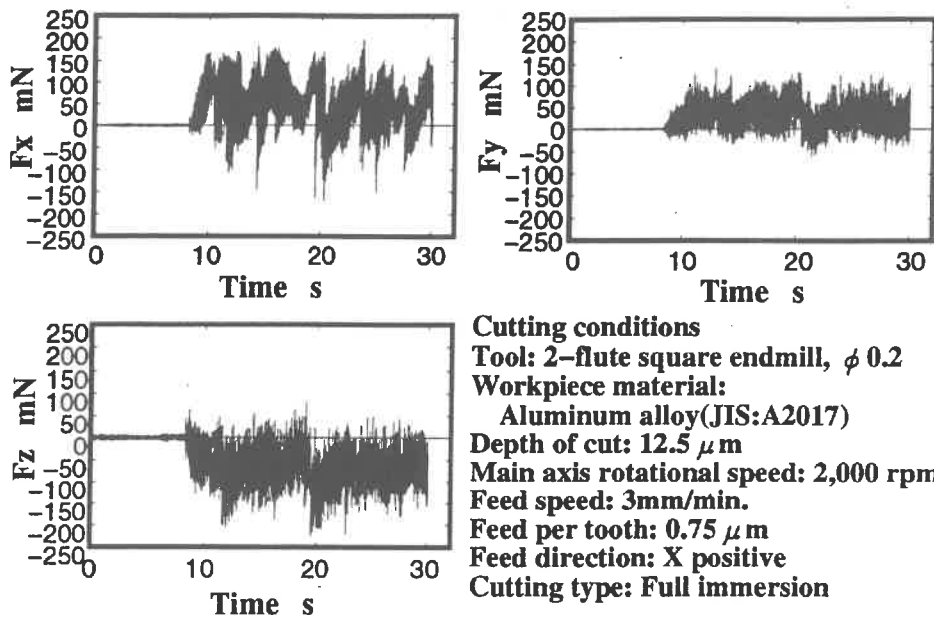


Fig.9 Experimental results using MMM

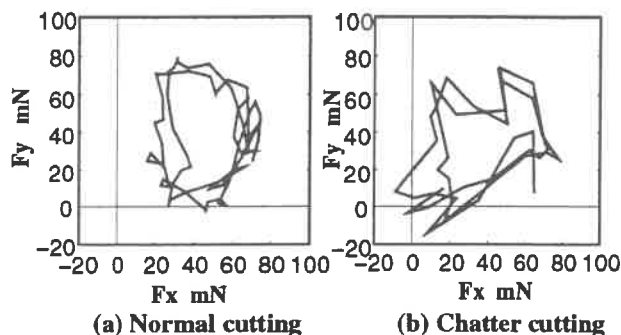


Fig.10 Force vector locus at micro cutting

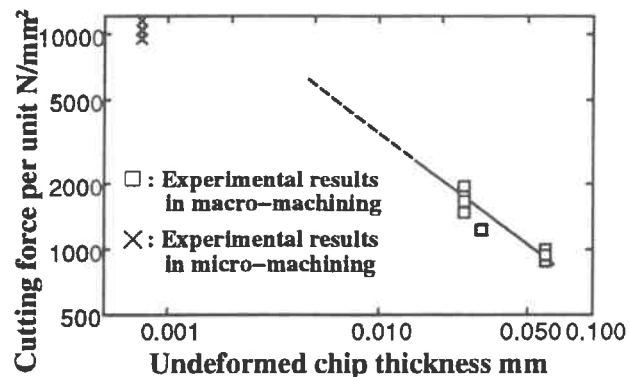


Fig.11 Relation between undeformed chip thickness and cutting force per unit

5. Conclusions

A micro-machining machine inside an SEM, with 3-axis force sensing capability and which can be operated from the remote place was developed. From the experiment, it is clear that the cutting force per unit increases as the cutting scale becomes smaller. Therefore, a high stiffness micro tool is necessary for micro cutting.

References

- [1] Hatamura, Y. et al., "Direct Coupling System Between Nanometer World and Human World," *Proc. IEEE Micro Electro Mechanical Systems*, pp.203-208, Napa Valley, California, 1990.
- [2] Johansson, S., "Hybrid Techniques in Micro-robotics," *Proc. IARP*, pp.72-83, Karlsruhe, 1993.
- [3] Takeuchi, Y., et al., "Ultraprecision 3D Micromachining of Glass," *Annals of the CIRP*, Vol.45/1, pp.401-404, 1996.
- [4] Friedrich, C., et al., "Mechanical Micromilling of PMMA," in *Proc. of the 10th Annual Meeting of the American Society for Precision Eng.*, pp.284-287, Austin, U.S.A., 1995.
- [5] Shimada, S., et al., "Molecular Dynamics Modeling of Thermal Conduction in Microcutting of Metals," in *Proc. of the 10th Annual Meeting of the American Society for Precision Eng.*, pp.21-24, Austin, U.S.A., 1995.
- [6] Shimizu, J. et al., "Research on Collision Mechanism of Atomic Cluster by Simulation," *Proc. Jap. Soc. Precision Eng. Spring Annu. Meeting*, pp.441-442, 1995 (in Japanese).
- [7] Tokushige, N., et al., "Observation of Cutting Behavior under Scanning Electron Microscope," *Proc. Jap. Soc. Precision Eng. Spring Annu. Meeting*, pp.897-898, 1995 (in Japanese).
- [8] Mitsuishi, M., et al., "Real-Time Machining State Detection Using Multiaxis Force Sensing," *Annals of the CIRP*, Vol.41/1, pp.505-508, 1992.

Ultraprecision Metal Cutting Process Monitoring Using Acoustic Emission -- Burr Formation

Jianshe Tang, Xuemei Chen and David Dornfeld
Department of Mechanical Engineering
University of California at Berkeley, Berkeley, CA 94720, USA

Abstract

The results of experimental studies of ultraprecision milling of aluminum are presented. The AE features correlated to burr formation in ultraprecision milling are discussed.

1. Introduction

Ultraprecision metal cutting (UPMC), in a contemporary sense, can be defined as a cutting technique which enables us to produce optical, mechanical and electronic components with micrometer or submicrometer form accuracy and a few tens nanometer surface roughness [Ikawa, 1991]. The accuracy of UPMC depends on the transfer characteristics of both machine tool motions and tool profiles into a workpiece. Submicrometer or even down to nanometer accuracies in machine tool motions and tool geometry are essential. Taking advantage of the development in advanced metrology, control technology and precision machine components, machine tools with motion accuracy at nanometric level and diamond tools with sharp cutting edge to as small as 20-45 nm were developed [Donaldson, 1986]. At present, ultraprecision metal cutting is extensively applied to the manufacturing of precision components, such as faceted optics used in laser systems, aluminum substrate drums in photo copying machine and aluminum substrates for computer hard disks.

There has been extensive investigations of UPMC in the area of chip formation and energy dissipation [Furukawa, 1988; Moriwaki, 1989; Ikawa, 1993; Lucca, 1994], where the cutting process is continuous and stable. Little research on the burr formation mechanisms and the in-process detection of burr formation in UPMC, where the cutting process is transient, has been done, even when the edge finishing is a critical specification for components such as polygon mirrors.

Acoustic emission is known to be sensitive to many process parameters and is evolving into an effective method for studying the process of material removal mechanisms and the condition of cutting tools in various metal cutting processes. The objective of this study is to investigate the connection between the burr formation and acoustic emission signal in ultraprecision milling.

2. Experimental approaches

2.1 Burr formation in face milling

During the milling operation, burrs are created where the milling cutter exits the machined part. Five different types of burrs can be formed with different combinations of milling conditions. It is generally accepted that both the height and thickness of burr are proportional to the undeformed chip thickness and usually are in the same order as the undeformed chip thickness [Nakayama, 1987; Chern and Dornfeld].

Figure 1 shows a diamond tool used in this test and the uncut chip geometry at two different depths of cut. The tool has a straight cutting edge and a circular cutting edge, with a transition edge between the straight and circular cutting edges. From the relationship between cutting parameters and tool geometry, it can be determined that when the depth of cut is smaller than 5 μm , see figure 1(a), work material will be removed only by the circular cutting edge of nose radius $R=100\text{mm}$. In this case, the maximum effective depth of cut can be calculated as $t_{\max} = f(2d/R)^{1/2}$ and is in the range of several to several tens of nanometer for our test conditions. The exit burr generated should be in the same order and is negligible. Hence, a burr free surface can be expected. However when the depth of cut is larger than 5 μm , the straight cutting edge will be involved in cutting, and the maximum effective depth of cut can be calculated by $t_{\max} = f \cos(\phi)$ and is on the same order as the feed rate. As feed rate is usually from 6 $\mu\text{m/r}$ to 40 $\mu\text{m/r}$ in our test setup, a surface with burrs with size in tens of micrometers is formed in this case. By monitoring and comparing the AE signals generated in these two cases, it is expected to see some AE features related to burr formation. The workpiece material in this experiment is as received 2024-T35 aluminum which is widely used for precision components like metal optics.

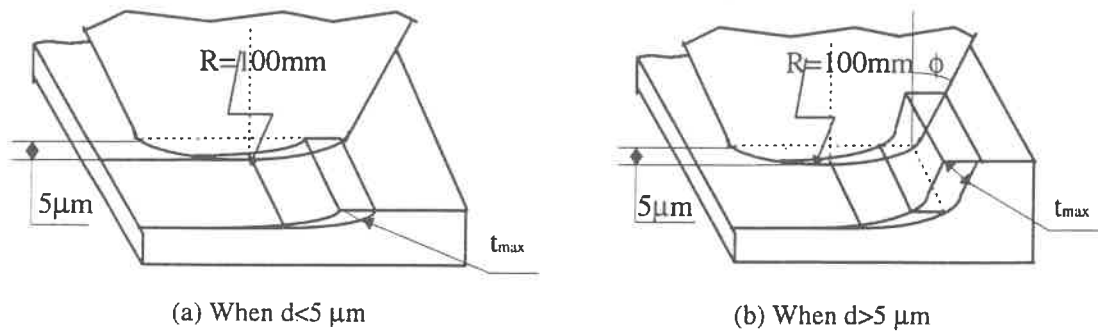


Fig. 1 Schematics of a diamond tool and chip geometry

2.2 AE sources in burr formation and AE measurement setup

A burr formation process can be divided into three stages, namely initiation, development and final burr formation [Ko and Dornfeld, 1991; Hashimura and Dornfeld, 1995]. The AE sources in both continuous cutting and burr formation are shown in figure 2. In the burr formation process, in addition to the deformation in primary, secondly and tertiary shear zone which are the dominant AE sources in continuous cutting, a negative shear zone and a crack propagation process is evident [Hashimura and Dornfeld, 1995]. These additional sources will constitute AE. Therefore, the AE energy in a burr formation process should be higher than that in continuous cutting and the difference in magnitude could be used for an indication of burr formation. As these regions vary in contribution to the AE signal with time (with tool motion), this can also be a feature for burr detection.

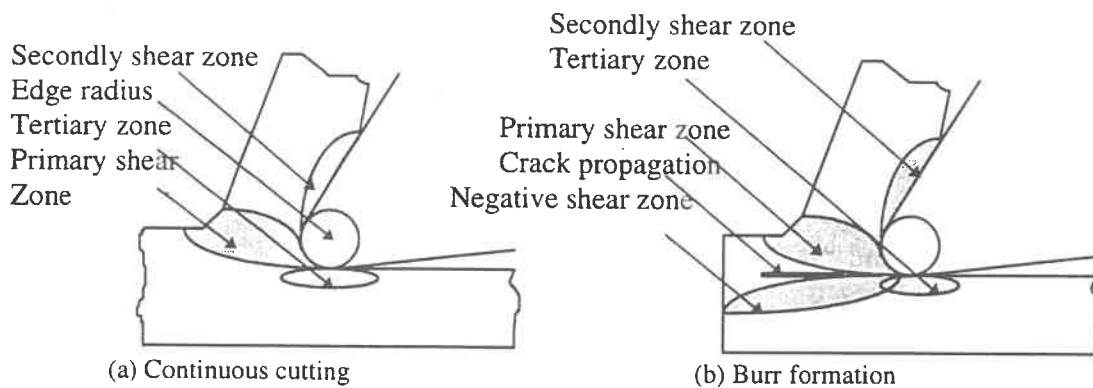


Fig. 2 AE sources in continuous cutting and burr formation process

A special arrangement developed to measure the acoustic emission is schematically illustrated in Figure 3. The workpiece was attached to the work table and a Dunegan model D9201A wide-band (50 KHz to 1 MHz) piezoelectric transducer mounted underneath the workpiece was used to measure the acoustic emission signal. The raw AE signal was fed through a preamplifier with fixed 40dB gain and built-in high-pass filter with 50 KHz corner frequency. This output was further amplified by a Dunegan amplifier with 20 dB gain and fed to an AE rms meter with sampling time of 0.25 ms and then digitized on-line through an A/D channel of the data acquisition board with commercial software GageScope on an IBM PC-AT for further analysis.

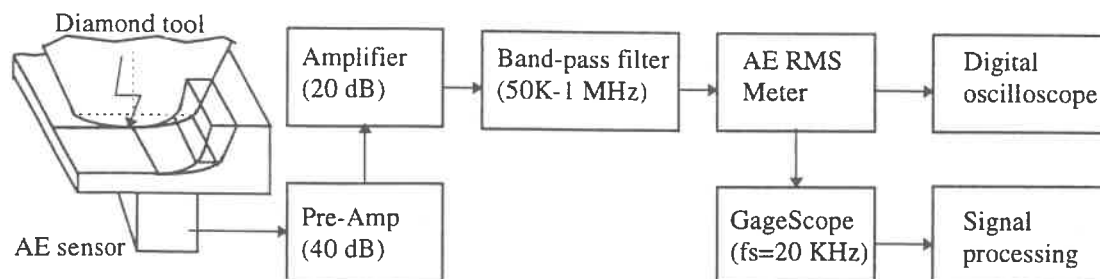


Fig. 3 Instrumentation for monitoring cutting process

3. Experimental results and discussions

Experiments were conducted on a Kuglar F380/1000 flycutting machine. The bed, guide block, X-slide and Z-column are manufactured from natural granite which gives excellent stiffness, thermal stability and good vibration damping of the machine structure. Machining spindle, X and Y axis slide tables, and work table are vacuum preloaded air-bearings. The infeed resolution in the vertical direction is $0.1\text{ }\mu\text{m}$. Figure 4 shows two surfaces generated at feed rate $f = 30\text{ }\mu\text{m/r}$, depth of cut $d = 0.5\text{ }\mu\text{m}$ and $20\text{ }\mu\text{m}$, that is $t_{\text{max}} = 0.09\text{ }\mu\text{m}$ and $15\text{ }\mu\text{m}$. As seen in the photos, surfaces with a burr and without a burr can be produced by changing the depth of cut d .

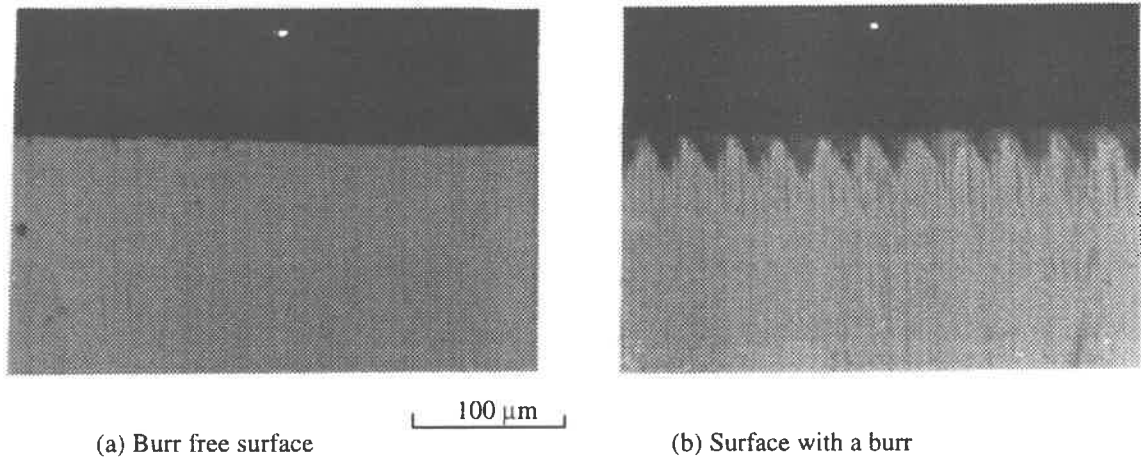


Fig. 4 Surfaces generated at different undeformed chip thickness

Figure 5 shows a chip generated at a depth of cut $d=0.1\text{ }\mu\text{m}$, and feed rate $f=6\text{ }\mu\text{m/r}$. This corresponds to a maximum uncut chip thickness $t_{\text{max}} = 8.5\text{ nm}$, and uncut chip width $C_w=144\text{ }\mu\text{m}$. It is clear from the photo that the machined chip width is about $50\text{ }\mu\text{m}$, which is only $1/3C_w$. This is because the uncut chip thickness is not constant along the cutting edge. It takes a maximum value at the top of the uncut shoulder and gradually changes to 0 at the bottom of cutting. There is a transition value below which no chip will be generated. With the photos, it can be estimated that in this machining setup, the minimum uncut chip thickness in which chip can be generated is about 6 nm . Fig. 6 is the corresponding AErms signal. The signal-to-noise ratio was on the order of 10^3 , which shows the very high sensitivity of AErms to work-tool contact even in the range of nanometer chip dimensions.

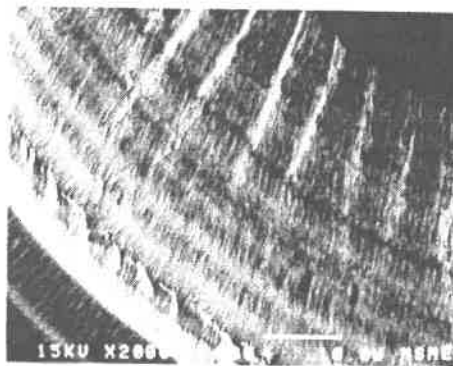


Fig. 5 A typical chip generated in UPMC

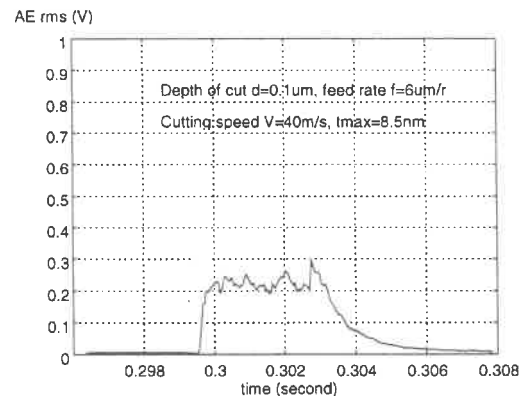


Fig. 6 Signal in burr free surface generation

Fig. 7 is the signal collected under the cutting conditions of $d=40\mu\text{m}$, $f=6\mu\text{m/r}$, that is, $t_{\text{max}} = 3\mu\text{m}$. The average AErms is higher and the pattern is obviously different from that in Fig. 6. The peak in the exit point is an indication of a burr formation. This is, in fact, similar to the AE signal observed by Diei and Dornfeld [Diei and Dornfeld, 1987] in conventional flycutting milling tests. Fig. 8 shows the exit peak AErms variation with depth of cut d at several feed rates. It can be found that the exit peak AErms is sensitive to both depth of cut and feed rate, this implies that AE is also sensitive to burr size variation. The transition from a burr free surface to a surface with a burr at about $5\mu\text{m}$ can also be clearly identified from the figure.

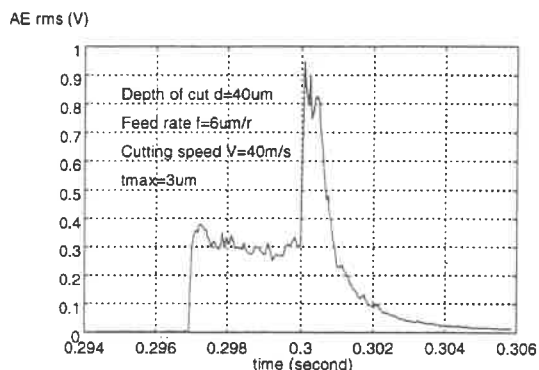


Fig. 7 Signal in burr formation

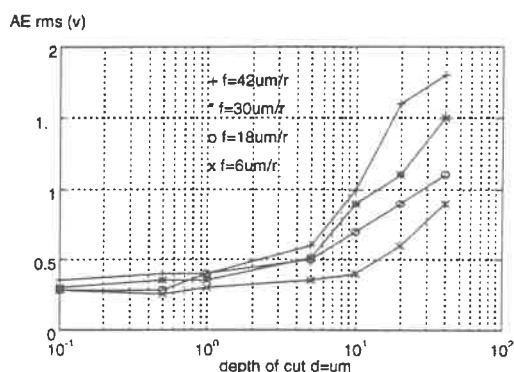


Fig. 8 AE rms peak value vs. cutting parameters

4. Conclusions

With a specially designed diamond tool, a burr free surface and a surface with burr can be generated by changing the depth of cut. The AE rms peak value at the exit point of cutting was found to be closely related to the burr formation from both AE source analysis and cutting tests. This AE feature could be used for in-process detection of burr formation.

References

1. Chern, G., et al., "Burr/breakout development and experimental verification", Trans. ASME, J. Eng. Mats. Tech., to appear.
2. Diei, E., et al., "Acoustic emission from the face milling process--the effects of process variables", Trans. ASME, Journal of Engineering for Industry, Vol. 109, 1987, pp92-99.
3. Donaldson, R., et al., "Design and performance of a small precision CNC turning machine", Annals of the CIRP, Vol. 36, No. 1, 1986, pp373
4. Furukawa, Y., et al., "Effect of material properties on ultraprecision cutting processes", Annals of the CIRP Vol. 37/1/1988, pp113-116.
5. Hashimura, M., et al., "Analysis of three dimensional burr formation in oblique cutting", Annals of the CIRP, Vol. 44/1/1995, pp23-26.
6. Ikawa, N., et al., "Ultraprecision metal cutting--The past, the present and the future", Annals of the CIRP, Vol. 40/2/1991, pp587-594.
7. Ikawa, N., et al., "Chip morphology and minimum thickness of cut in micromachining", Journal of the Japan Society of Precision Engineering, Vol. 59, No. 4, 1993, pp141-147, (in Japanese).
8. Ko, S. and Dornfeld, D., "A study on burr formation mechanisms", Trans. ASME, J. Eng. Mats. Tech., Vol. 113, No. 1, pp 75-87.
9. Lucca, D., et al., "Aspect of surface generation in orthogonal ultraprecision machining", Annals of the CIRP Vol. 43/1/1994, pp43-46.
10. Moriwaki, K., et al., "Machinability of copper in ultraprecision micro diamond cutting", Annals of the CIRP Vol. 37/1/1989, pp115-118.
11. Nakayama, K., et al., "Burr formation in metal cutting", Annals of the CIRP, Vol. 36/1/1987, pp33-36.

Controlled Microcutting Tests on Single Crystal Silicon

B.V. Tanikella and R.O. Scattergood

Precision Engineering Center
Department of Materials Science and Engineering
North Carolina State University, Raleigh, NC 27695-7918

INTRODUCTION

Precision grinding [1] is accomplished by the cutting action of many sharp diamonds embedded in the wheel binder material. The fracture damage in the resulting part is dependent on the material properties and operating conditions. Considering grinding as a multi-scratch process, the response of a single cutting grain and the damage transitions produced by the cutting action as the load on the grain is varied, will be of primary interest. The microcutting action of a single cutter is important because it plays a key role in determining the subsurface fracture damage produced in processes such as polishing and grinding [1,2]. In recent studies with glass and amorphous SiC thin films, Tanikella and Scattergood [3,4,5] demonstrated a one-to-one correspondence between crack-initiation and depth-of-cut during microcutting. Extension of the results using statistical averaging effects will facilitate the prediction of the surface finish and subsurface fracture damage. The effect of crystallography is important for developing these models for crystalline materials. The purpose of the present work is to investigate the cracking thresholds, crystallography, and tool geometry effects for single crystal silicon. An understanding of the micromechanics of silicon would also enhance its utility as a high-precision, high-strength, high-reliability mechanical material, especially applicable in micro mechanical devices and components in conjunction with integrated electronics.

EXPERIMENTAL DETAILS

The material used in the present study was device grade single crystal p-type silicon with (100) and (111) surfaces. A freshly sharpened Vickers diamond pyramid indenter was used as the cutting tool for a series of tests and its condition was monitored frequently. All of the tests were conducted in controlled laboratory air at $20 \pm 0.05^\circ \text{C}$ and $40 \pm 10\%$ relative humidity. A field emission JOEL 6400 scanning electron microscope (SEM) was used for micrographic observations. Two different test systems [4,5,6] were used for conducting the microcutting tests and they adopt either force control or depth control.

In the force controlled setup [4], a $140 \mu\text{m}$ range piezo translator (PZT) was mounted to the moveable crosshead of an ATS benchtop load frame and the PZT amplifier was interfaced to a PC for closed-loop load control profiles during microcutting tests. A 10 N load cell and fixture to hold a Vickers diamond indenter were attached to the PZT. An x-y stage was mounted to the base of the load frame for sample positioning. One axis of the stage was driven by a motorized micrometer screw using an analog encoder-controller for speed control. An acoustic emission sensor with a frequency-detection window set between 100 KHz to 900 KHz (200 KHz resonance frequency) was used for the acoustic measurements. A nominal cutting speed of $5 \mu\text{m}/\text{sec}$ was used for all of the microcutting tests. The microcutting grooves for different maximum loads appear as shown in Figure 1a. The length of the entrance, exit and hold portion of each groove is constant because of the fixed cutting speed, but the groove widths change with maximum load as indicated. During a loading profile the acoustic pulses detected by the oscilloscope were transferred and stored in the PC (using Labview® software) for comparison with the load values.

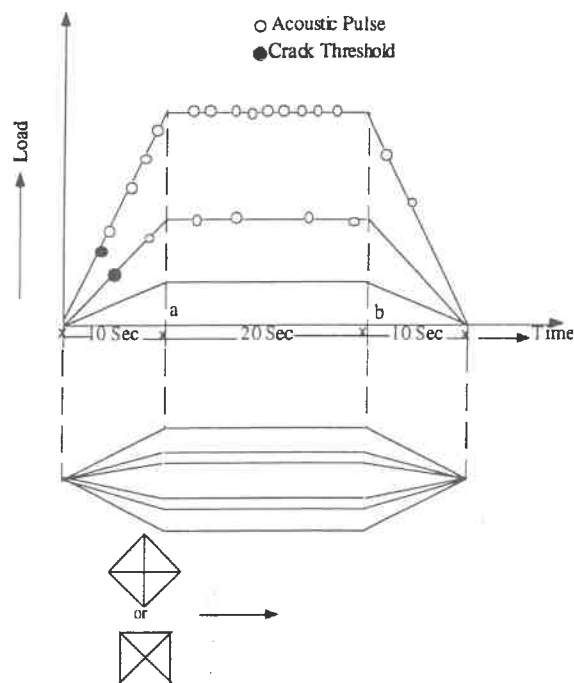


Figure 1a.

Test scheme during force control Tests

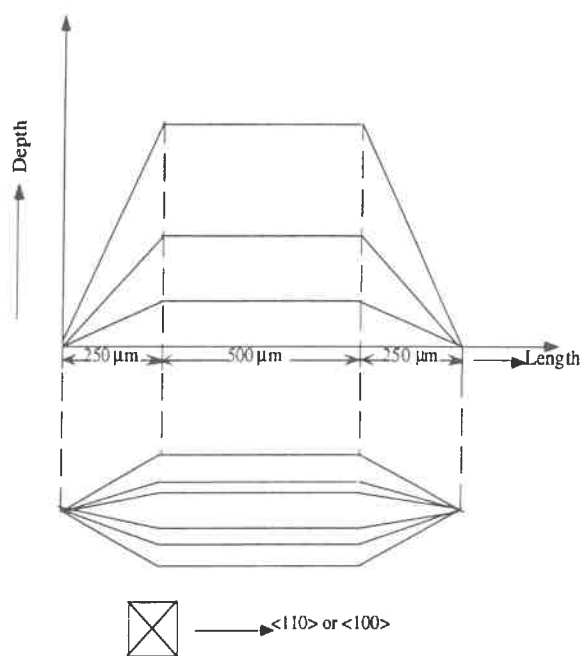


Figure 1b

Test scheme during depth control Tests

In the depth controlled setup [6], a Rank Pneumo ASG 2500 diamond turning machine (DTM) was used for conducting microcutting tests. This DTM features T-base configured hydrostatic bearing slides, an aerostatic bearing workpiece spindle, and laser interferometer feed back. The spindle of the diamond turning machine was locked mechanically and the vacuum chuck mounted on it was used to hold the sample. The sample was a 4 inch wafer and was mounted on the vacuum chuck of the DTM in as-received condition. $\langle 110 \rangle$ and $\langle 100 \rangle$ crystallographic directions of the wafer were adopted as two different cutting directions, and the direction was fixed before locking the spindle. A three-axis PZT load cell was used for recording the force data. The load cell and a specially fabricated fixture were mounted on a custom built dynamometer. Viscous grease was used between the interfaces and the assembly was held together with a torque of 11.3 N-m, to achieve a stiff setup. The slides of the DTM were programmed to perform the microcutting tests. A schematic of the test scheme is shown in Figure 1b. The x-slide of the DTM was programmed to give a nominal speed of 10 mm/min during all the microcutting tests. The side of the Vickers diamond is along the cutting direction. In earlier studies using the force controlled setup, it was found out that cutting along the side of the indenter produces less fracture damage compared to cutting along the diagonal for silicon. Performing constant depth-of-cut tests on single crystal silicon is complicated due to the nonflatness in silicon wafers due to residual stresses. A capacitance gage was used to determine the slope before conducting each microcutting test and the data thus obtained was applied in the part program assuming a linear slope. Thus, the resulting microcutting grooves would have the desired depth at a given position along the microcutting groove. The force data was acquired into a PC using Labview® software.

RESULTS AND DISCUSSION

Force Control

Microcutting tests as shown in Figure 1a were conducted in the load range of 0.03 to 3 N. The variables during microcutting were load, indenter orientation, cutting direction in relation to the crystal direction, and the normal to the crystal surface (i.e. (100) or (111)). The crack-initiation threshold, as manifest by the detection of an acoustic pulse during the load cycle in Figure 1a, was 0.04 to 0.4 N, depending upon test conditions. The threshold observed during the loading part of

the cycle varied significantly with the loading rate. Figure 2 shows the effect of loading rate while cutting along the diagonal on a (100) surface along one of the $\langle 110 \rangle$ directions. Similar loading rate effects were observed for various cutting direction combinations. Each data point in Figure 2 shows the first acoustic pulse observed during the load cycle, indicating that the crack-initiation threshold during loading can be increased with increasing loading rate. However, the amount of increase depended somewhat upon the direction of cutting and the crystal surface plane.

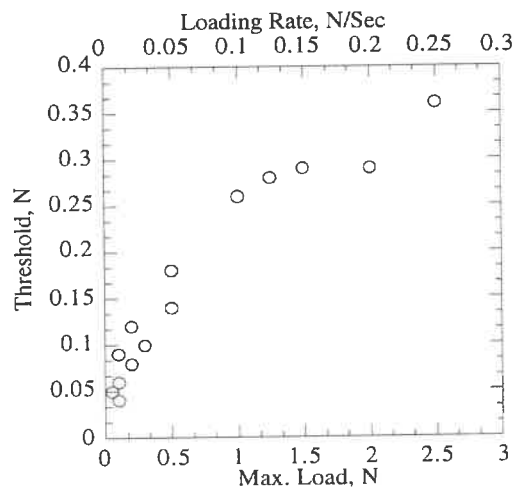


Figure 2
Plot showing the effect of loading rate

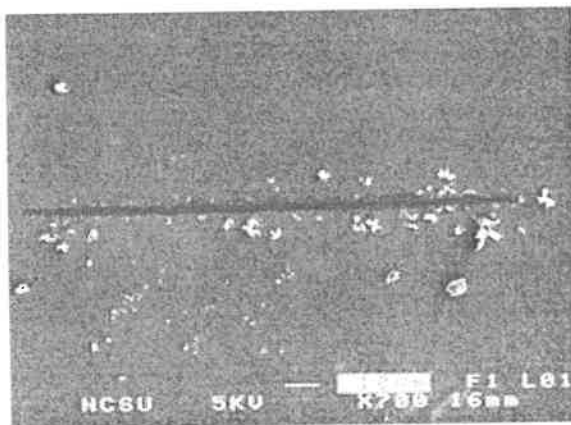


Figure 3a
Microcutting groove at a load of 0.04 N

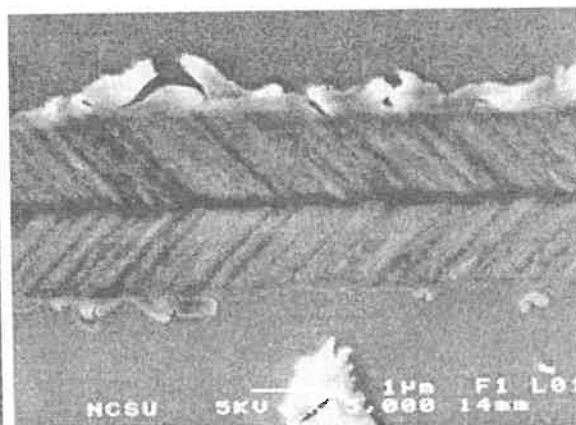


Figure 3b
A close-up look at the groove Figure 3a

The extent of fracture damage in the microcutting grooves depended strongly upon the cutting direction and the indenter orientation, in relation to the crystallographic orientation of the wafer. Figure 3a and 3b show a microcutting groove at the threshold 0.04 N. Detailed SEM observations showed that the groove is crack-free and is a ductile cut. Figure 4 shows a similar groove as in Figure 3b with the load just above the threshold, at 0.05 N. A tiny radial crack along the edge of the groove can be observed. This is the kind of cracking observed just above the threshold. Note that Figures 3b and 4 have opposite cutting directions. At the threshold, the cracking behavior was somewhat erratic and cracking events were not reproducible. As the load is increased the extent of damage increased, however, and it depended strongly upon the crystallography and the indenter

orientation in Figure 1a. The indenter oriented with its diagonal along the cutting direction produced more damage compared to the cutting grooves made with the indenter side along the cutting direction. This is in contrast to the observations with glassy amorphous materials such as borosilicate glass [4,5], wherein the extent of damage produced is less while cutting along the diagonal of the tool. Thus, in the case of single crystal silicon, a wedge-shaped tool produces more fracture damage compared to a flat-faced tool.

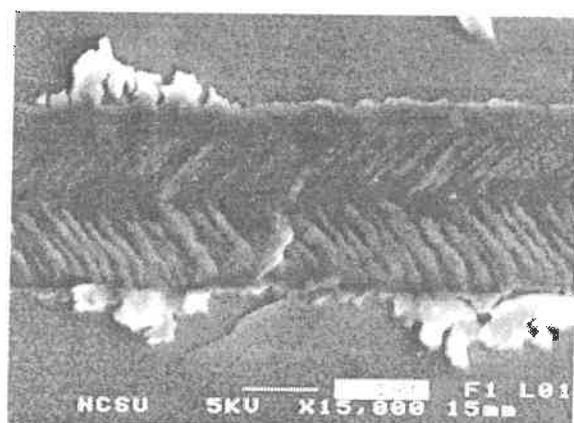


Figure 4

A close-up look at a 0.05 N microcutting groove

The effect of crystallography along with the indenter orientation provides an optimal direction for cutting. There were four different configurations when working with a (100) wafer. These constitute cutting along $\langle 110 \rangle$ and $\langle 100 \rangle$ directions, with two different indenter orientations. Cutting along $\langle 100 \rangle$ direction with the side of the indenter along the cutting direction is the optimal combination for minimal fracture damage in the microcutting grooves. Figures 5a and 5b show comparison of two microcutting grooves (with crystal surface (100)) at a load of 0.5 N, with differences in the indenter and crystallographic orientations. Figure 5a had the cutting direction along $\langle 100 \rangle$ with the indenter side oriented along the cutting direction, while Figure 5b had the cutting direction along $\langle 110 \rangle$ with the indenter diagonal oriented along the cutting direction. Figure 5a shows significant portions of ductile areas, while the groove in Figures 5b shows severe chipping damage. The crystallographic relations are in agreement with the observations in literature [7]. Similar crystallographic effects were observed in the case of (111) wafers [8].

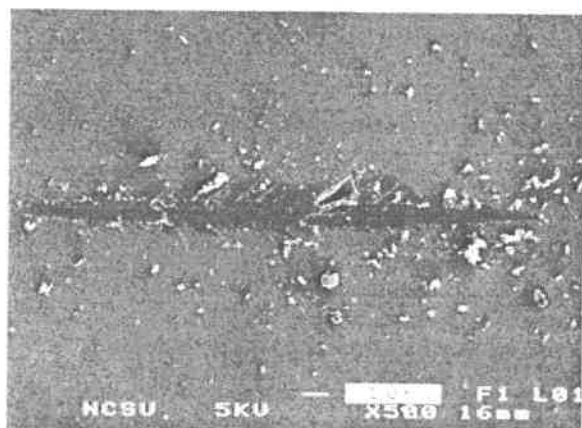


Figure 5a

Microcutting groove at 0.5 N along $\langle 100 \rangle$ cut with the side of the Vickers indenter.

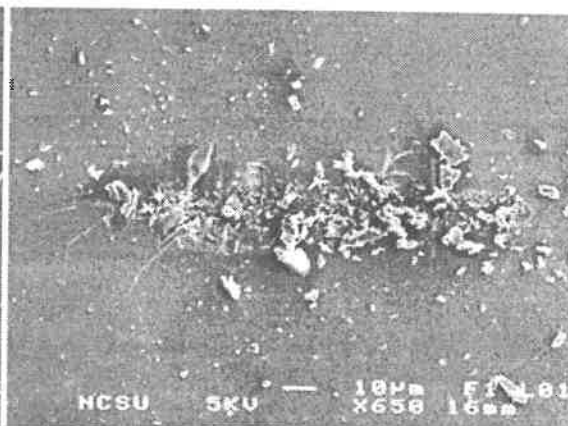


Figure 5b

Microcutting groove at 0.5 N along $\langle 110 \rangle$ cut with the diagonal of the Vickers indenter.

Depth Control

Figure 6a shows a plot of normal force vs time for a microcutting groove at a depth-of-cut of 200 nm. Figure 6b shows an SEM micrograph corresponding to a portion of the groove in Figure 4, at the maximum cutting depth. Closer examination showed that the whole groove is crack-free and ductile. Also, the force is constant (above the noise level) at the maximum depth, indicating a uniform depth-of-cut. If the correction for the surface tilt of the wafer was inappropriate, the force profile would show a tilt of the force trace at the maximum depth-of-cut, indicating the deviation from a constant depth-of-cut.

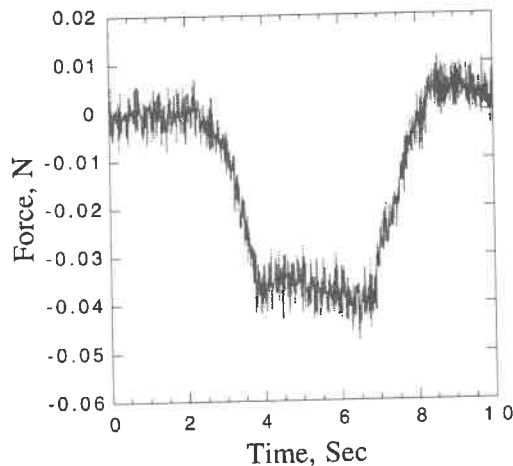


Figure 6a
Normal Force vs Time for a 200 nm depth-of-cut microcutting groove

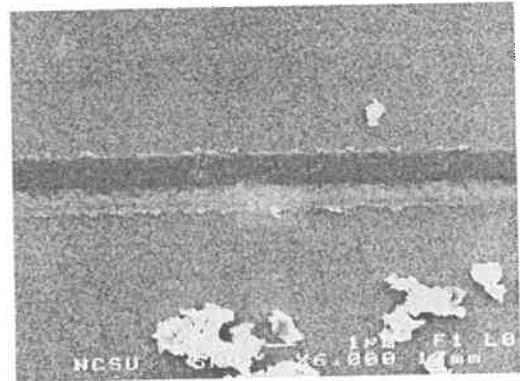


Figure 6b
SEM micrograph of the microcutting groove corresponding to Figure 6a at maximum depth

Above the cracking threshold, the force plots show significant deviation from Figure 6. Cracking thresholds are manifest as (unloading) pulses in the force traces. The threshold can be expressed in terms of a critical force or a critical depth-of-cut, since the tool geometry is known. As one would expect, results obtained from the depth-control tests match very well with the cracking thresholds obtained using the force control system. Figure 7a shows the force profile corresponding to a microcutting groove just above the threshold, at a depth-of-cut of 250 nm. In Figure 6a, the maximum normal force could be observed to be below the threshold, and the groove was of ductile nature. The maximum force in Figure 7a is about 0.06 N, and this is clearly above the threshold.

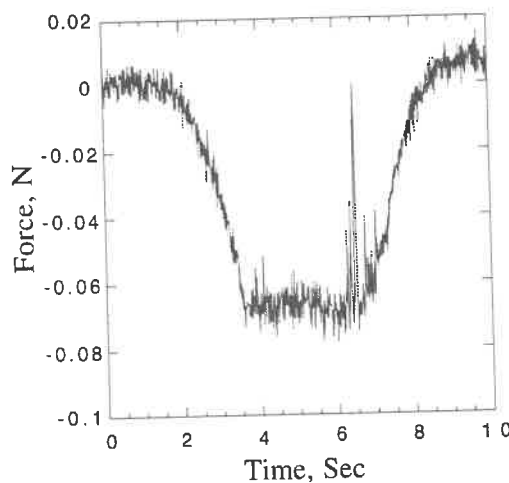


Figure 7a
Normal Force Profile for a 250 nm depth-of-cut microcutting groove

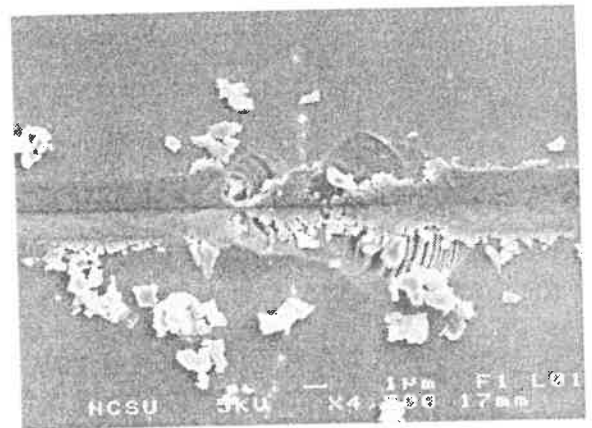


Figure 7b
SEM micrograph showing the fracture event associated with the large force pulse in Figure 7a

Detailed SEM observations showed a one-to-one correspondence between these force pulses in Figure 7a and a fracture event within the groove. The lateral chip in a portion of the microcutting groove shown in Figure 7b was associated with the largest force pulse in Figure 7a. This large force pulse indicates that the tool is unloaded to a large extent due to this large chip (or missing material ahead of the tool tip), and the force comes back to the normal position upon subsequent cutting to maintain the programmed depth. Similarly, the smaller force pulses corresponded to radial cracking and smaller scale chipping damage.

Figure 8a shows the normal and cutting force vs time respectively, for a microcutting groove at a programmed depth-of-cut of 1000 nm. It is evident that the depth-of-cut for this groove is well above the critical depth-of-cut. The initial force pulse during the load cycle occurs well above 0.045 N, indicating an increase in the cracking threshold as the loading rate goes up (loading rate increases with the depth-of-cut because the ramp distance in Figure 1b is fixed and the cutting speed is fixed). The initial pulse in Figure 8a occurs at a load close to 0.12 N. SEM observations showed no evidence of fracture in the microcutting groove before this initiation point. The loading rate effect clearly supports the results in Figure 2. The number of force pulses in Figure 8a far exceeds the number in Figure 7a. The cutting force profiles also showed a force pulse corresponding to each large pulse on the normal force profile. This indicates that there is also a significant force change in the cutting direction due to the chipping action. Detailed SEM observations of the microcutting groove corresponding to Figure 8a showed a one-to-one correspondence between the fracture events and the force pulses. Qualitatively, the magnitude of the force pulses were proportional to the extent of fracture. No attempts were made to provide quantitative measure of the extent of fracture. Figure 8a clearly shows that the tool is completely unloaded at the largest force pulse and this corresponded to the SEM micrograph shown in Figure 8b. There is a lateral chip ahead of the tool, and the tool is in a completely unloaded state for a short period due to the missing material, as indicated by the force information.

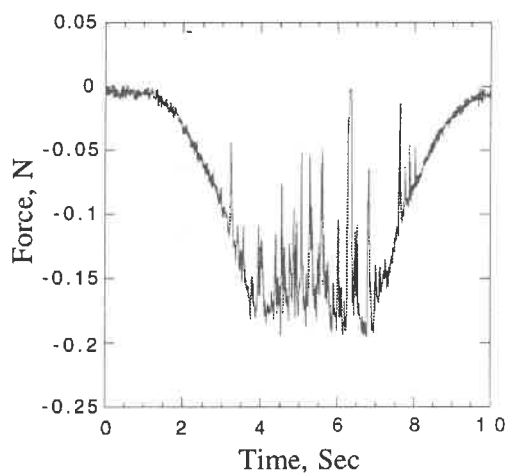


Figure 8a

Normal Force vs Time for a 1000 nm depth-of-cut microcutting groove

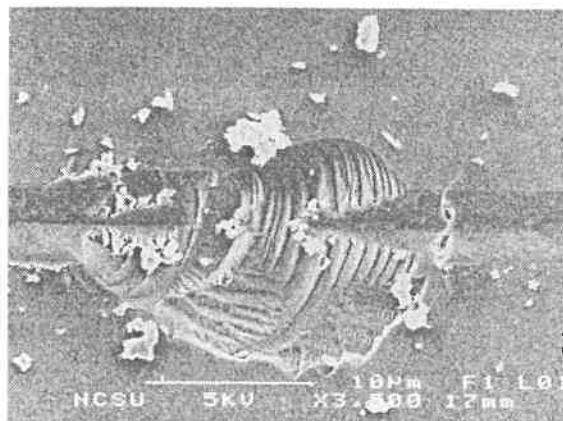


Figure 8b

SEM micrograph showing the fracture event associated with the large force pulse in Figure 8a

All the above results were obtained for cutting along the $\langle 110 \rangle$ direction on the (100) silicon wafers. Changing the cutting direction to $\langle 100 \rangle$ (by rotating the spindle by 45°), showed significant differences in the fracture morphologies of the microcutting grooves. Figure 9a shows the normal force corresponding to microcutting groove at a depth-of-cut of 1200 nm. This force curve doesn't show any force pulses similar to those in Figure 8a. The maximum normal force in this case is about 0.2 N, and is higher than that in Figure 8a. Even at such high force levels, there was no indication of large scale fracture associated with the present cutting groove. Detailed SEM

observations showed no lateral chips along the whole groove. However, the groove was observed to have radial cracks along both sides of the groove as shown in Figure 9b. This can be explained from consideration of the tool geometry, cleavage planes and directions. Indentation fracture [9] shows that the radial cracks originate from the corners of a Vickers indenter. When the cutting direction is along $\langle 100 \rangle$, the radial cracks on $\{111\}$ cleavage planes would initiate in the $\langle 110 \rangle$ directions. Since the $\langle 110 \rangle$ directions are the favorable directions for crack propagation, the radial cracks extend and show up along the edges of the cutting grooves. While cutting along $\langle 110 \rangle$, the radial cracks could not propagate along the $\langle 100 \rangle$ directions and thus causing significant residual stress build-up. The residual stresses thus cause large scale chipping rather than radial cracking, for this geometry.

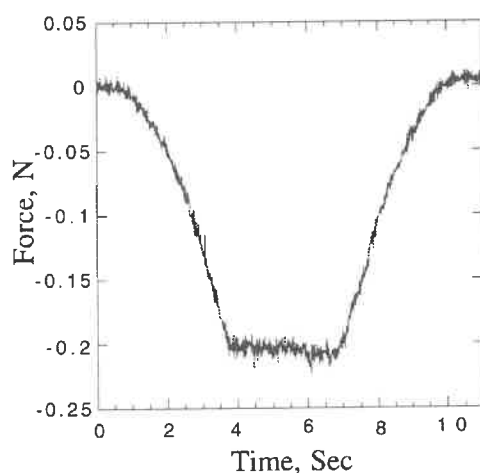


Figure 9a

Normal Force vs Time for a 1200 nm depth-of-cut microcutting groove, while cutting along $\langle 100 \rangle$

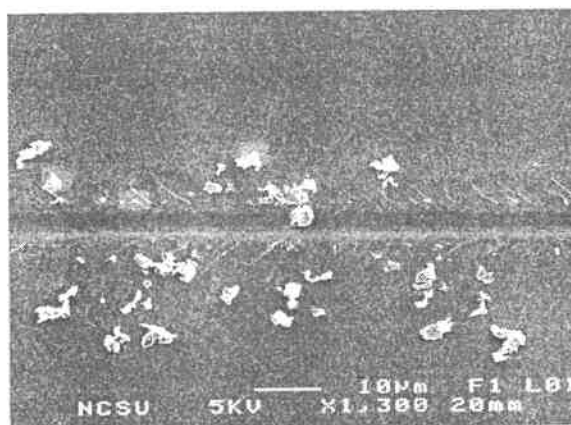


Figure 9b

SEM micrograph showing typical fracture morphology while cutting along $\langle 100 \rangle$ direction

Microplasticity

The features in the grooves and the debris around the microcutting grooves suggested pressure induced phase transformation during microcutting [10]. MicroRaman analysis confirmed the presence of amorphous Si phase within microcutting grooves and particle debris removed from the grooves. This can be attributed to the formation of the β -tin metallic Si phase during the cutting process. Periodic features observed within the grooves (Figures 3b and 4) are consistent with an intermittent shear-type of chip formation process that will be facilitated by the formation of a ductile metallic phase. The MicroRaman scans also indicate the presence of large tensile residual stresses that are attributed to a mixture of amorphous and crystalline phases produced by microcutting. In situ recrystallization of the amorphous phase during microcutting has been suggested as a mechanism [10].

SUMMARY

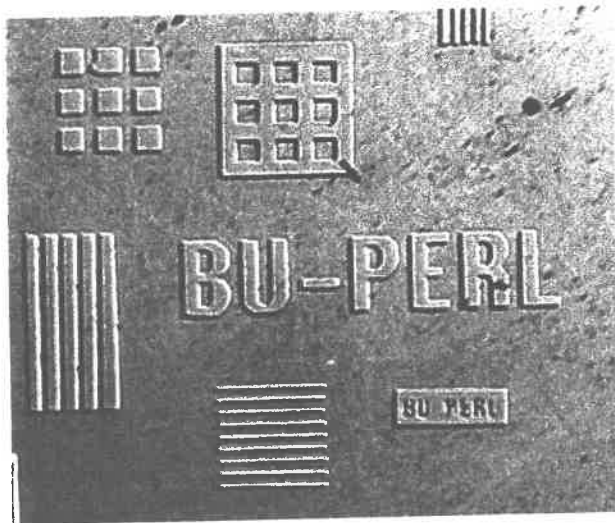
Crack-initiation studies were conducted using two different test systems using force control and depth control respectively. Acoustic emission and force sensing were used as the tools for determining the cracking thresholds. An increase in the cracking threshold with loading rate was observed, and this would have important ramifications with respect to diamond turning and grinding. Cutting along the side of the Vickers indenter and $\langle 100 \rangle$ directions of the wafer were found to produce minimal fracture damage. Pressure induced phase transformation of silicon was understood to be the reason for periodic features in the grooves.

REFERENCES

1. T. Bifano, T. A. Dow, and R. O. Scattergood, "Ductile-Regime Grinding: A New Technology for Machining Brittle Materials", *J. of Engr. for Ind.*, **113**, pp 184-189, 1991.
2. P. Blake et al., "Precision Machining of Brittle Materials", *Ceramic Bulletin*, **67** [6], (1988).
3. B.V. Tanikella and R.O. Scattergood, "Acoustic Emission during Indentation Fracture", *J. Am. Ceram. Soc.*, **78** [6] pp 1698-702, 1995.
4. B.V. Tanikella and R.O. Scattergood, "Fracture damage in Borosilcate Glass During Microcutting Tests" *Scripta Met.*, **34** [2], pp 207-213, 1996.
5. B.V. Tanikella, K.A. Gruss, R.F. Davis, and R.O. Scattergood, "Indentation and Microcutting Fracture Damage in a SiC coating on an Incoloy Substrate" *surf. coat. tech.*, in press.
6. B.V. Tanikella and R.O. Scattergood, "Controlled Microcutting Tests on Semiconductor Silicon Crystals using a Diamond Turning Machine", *Prec. Engg.*, to be submitted.
7. W.S. Blackley and R.O. Scattergood, "Crystal Orientation Dependence of Machining Damage-A Stress Model", *J. Am. Ceram. Soc.*, **73** [10], pp 3113-15, 1990.
8. B.V. Tanikella, PhD thesis, North Carolina State University, in progress, 1996.
9. B.R. Lawn, "Fracture of Brittle Solids - second edition", Cambridge University Press, 1993.
10. B.V. Tanikella, A.H. Somasekhar, A.T. Sowers, R.J. Nemanich, and R.O. Scattergood, "Phase Transformation During Microcutting Tests on Silicon", *Appl. Phys. Let.*, to be submitted.

Patterning Submicrometer Features Using Masked, Broad Beam Ion Machining

Helen E. Caggiano and Thomas G. Bifano
Boston University Precision Engineering Research Laboratory
Boston, MA



Broad beam ion machining is being investigated as a process for imprinting submicrometer features on advanced ceramic and glass materials. Ion machining has been used previously for final surface contour correction of mirrors [1 - 5]. This paper describes using broad, neutral ion beams to create patterns on the surface of a masked substrate. Photolithography techniques are used to mask the substrate with photoresist prior to ion machining. Portions of the substrate beneath the photoresist mask are shielded from the ion beam while the unmasked portions of the substrate are machined. While the substrate is being machined, the mask is also being machined.

To date, our research indicates that the achievable lateral resolution of this technique is limited only by the resolution of the photoresist (approximately 50 nm). Vertical precision is on the order of 10 nm. Several combinations of photoresist and substrate have been evaluated for various applications (i.e. diffraction gratings, video disc masters, compact disc masters, and magneto-optical storage devices). A calibration mask has been produced to evaluate achievable pattern resolution in the photoresist layer. After ion machining, the quality of feature transfer to the substrate has also been evaluated. It has been observed that when subjected to the intense energy of ion machining, the photoresist may flow and lose edge definition. Heat transfer in the ion machining process has been measured and numerically modeled. An ion machining system (IMS) has been designed and constructed to conduct broad beam ion machining experiments on large samples using a 110 mm diameter ion source. Submicrometer features have been produced in various ceramics (SiC , Al_2O_3 , ZrO_2), glasses (fused silica, Corning 9600), metals (nickel, molybdenum), and semiconductors (silicon).

The IMS consists of an Ion Tech 110 mm ion source with a plasma bridge neutralizer (PBN). The source is mounted vertically in the IMS chamber. Samples are mounted directly beneath the source on a rotary stage.

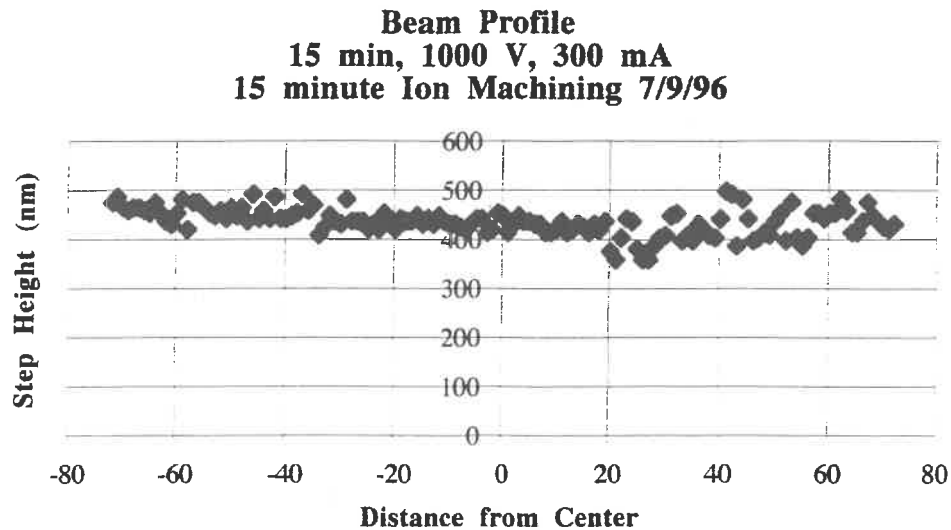


Figure 1. Beam profile of 110 mm ion source.

The ion source provides a nearly uniform beam over a 140 mm diameter. This allows large samples patterned with photoresist masks to be uniformly machined by the ion beam without introducing a rastering process.

As a result of this broad beam machining, the amount of heat energy directed toward the substrate is quite large. As a result, if the substrate temperature exceeds the melting temperature of the photoresist mask, the mask will flow, resulting in a loss of feature integrity. The heat transfer process has been investigated using silicon. An exposed junction thermocouple was placed under the substrate, using vacuum grease as a thermal conductor between the substrate and the thermocouple. The temperature of the substrate was measured for three different ion beam configurations as shown in Table 1.

Table 1. Ion source configurations for thermal testing.

Thermal Test	Beam Current (mA)	Beam Voltage (V)
1	300	1000
2	150	1000
3	140	500

Machining rate increases with increasing beam current and with increasing beam voltage. The measured temperature exhibit typical behavior of a first-order thermal system, as shown in Figure 2.

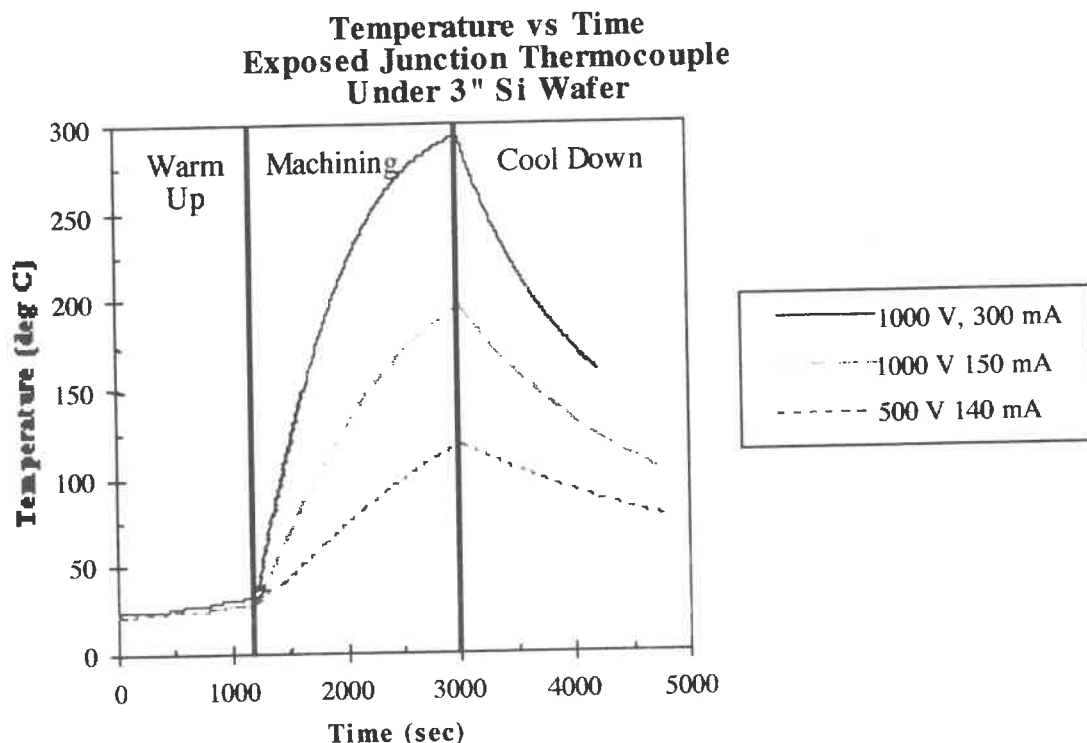


Figure 2. Measured temperature versus time for three machining conditions.

The curves in Figure 2 indicate the temperature measured during ion machining tests, during ion beam warm up, machining, and cooling. The first 1200 seconds (20 minutes) correspond to the warm up period during which the source is on (i.e. the ion source is generating an ionized plasma, but the ions are not being accelerated toward the substrate), and the plasma is stabilizing. The next 1800 seconds (30 minutes) correspond to machining, during which the ions are accelerated towards the substrate. The remaining 1200 seconds (20 minutes) indicates the cooling period, where the beam and source are shut down, and argon gas continues to flow. This cooling period is required to avoid oxidation of the source's tungsten filament as the chamber is vented to atmosphere. Argon continues to flow through both the source and the PBN until the chamber has cooled.

An analytical model of the heat transfer during these processes is under development. The model will be used to predict (and avoid) the onset of feature degradation during machining. The model can also be used to determine if auxiliary cooling systems are required and the amount of cooling needed during the machining process.

The machining rates for silicon were tested at each of these ion source configurations shown in Table 1, and the results are summarized in Table 2.

Table 2. Neutral ion machining rates for silicon using thermal testing configuration

Beam Voltage (V)	Beam Current (mA)	Silicon Material Removal Rate (Å/min)
1000	300	300
1000	150	150

Once the machining rate for silicon was determined, the samples were patterned using a photolithography mask. The patterns consisted of line widths from $0.3\ \mu\text{m}$ to as much as $20\ \mu\text{m}$, with features similar to those needed for compact discs, diffraction gratings, and magneto-optical storage devices. Four different types of negative-tone photoresists were used in the ion machining study to determine which photoresist maintained the best feature integrity. The features were observed using the SEM prior to and following the machining process. The following SEM photographs show the pattern transfer from the photoresist mask to an ion machined silicon wafer.

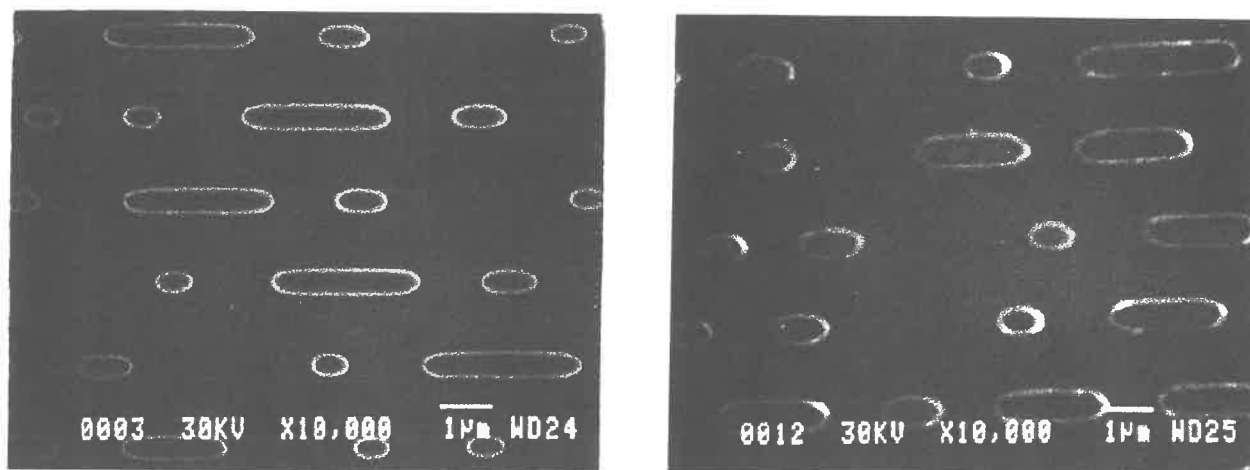


Figure 3. a. Photoresist mask on silicon wafer, prior to ion machining.

b. Patterned silicon wafer after ion machining.

From the photographs in Figure 3, direct transfer from the exposure to the machining processes can be observed. Other substrate materials will be tested using similar conditions to those used with the silicon wafers. From these experiments, alternative substrates can be developed for potential use as stampers for compact disc and video disc replication and substrates for diffraction gratings and magneto-optical applications. Once alternative substrates are determined, their utility in these different applications can be verified.

References:

- [1] Allen, Lynn N., Hannon, John J., and Wambach, Richard W., "Final surface error correction of an off-axis aspheric petal by ion figuring." *SPIE: Active and Adaptive Optical Components*, Vol. 1543, 1991, p. 190 - 200.
- [2] Allen, Lynn N. and Keim, Robert E., "An ion figuring system for large optic fabrication." *SPIE: Current Developments in Optical Engineering and Commercial Optics*, Vol. 1168, 1989, p. 33 - 50.
- [3] Allen, Lynn N., Keim, Robert E., and Lewis, Timothy S., "Surface error correction of a Keck 10 m telescope primary mirror segment by ion figuring" *SPIE: Advanced Optical Manufacturing and Testing II*, Vol. 1531, 1991, p. 195 - 204.
- [4] Drueding, Thomas W., Bifano, Thomas G., and Fawcett, Steven C., "Contouring Algorithm for Ion Figuring." *Precision Engineering*, Vol. 17, 1995. p. 10 - 21.
- [5] Fawcett, Steven C., Drueding, Thomas W., and Bifano, Thomas G., "Neutral ion figuring of chemical vapor deposited SiC." *Optical Engineering*, Vol. 33. No. 3. March 1994. p. 967 - 974.

Nanogrinding

Hans H. Gatzert and J. Chris Maetzig
Institute for Microtechnology (imt), Hanover University
Hanover, Germany

Abstract—Nanogrinding is an ultraprecision machining process that is based on a lapping kinematic and is very well suited for surface machining of advanced ceramics. While lapping uses loose grain, for nanogrinding the abrasive grain is completely embedded in a soft metallic plate which becomes the grinding tool. Therefore, nanogrinding is a two step process. The first step is to create the grinding plate, in a second step the work piece is machined. Nanogrinding requires a rough grinding plate, the roughness is created by conditioning the plate with pumice. During this process, pumice is embedded in the grinding plate and remains there throughout the process. Between pumice particles, the basic soft metallic plate material forms plateaus, in which diamond grain is embedded by a conditioning ring, which also aligns their summits coplanar to the plate surface. This arrangement allows to perform ductile machining, resulting in little material damage and excellent surface finish.

I. INTRODUCTION

The recording head industry is one of the dominant users of advanced ceramics, (like) as there are hot isostatic pressed (HIP) alumina-titanium carbide (70% Al_2O_3 , 30% TiC), calcium titanate (Ca TiO_3), and, at an experimental level, silicon carbide (SiC). Furthermore, to a small extend, the semiconductor material silicon (Si) is utilized, which has machining properties similar to ceramics. One of the machining challenges when manufacturing recording heads for rigid disk drives is to generate surfaces that are plane (or very close to plane, with a crown of some hundred nanometers per mm work piece length) and smooth (with an average roughness in the 1 nm range) [1, 2]. Both conditions are essential for achieving good tribological behavior. Due to the fact that such a recording head is gliding on a diamond like carbon (DLC) coated disk for part of its operation, the called for excellent tribological behavior is mandatory.

The two dominant processes in machining flat surface planes are grinding and lapping [3, 4]. The atomic bond of ceramic materials is covalent, ionic, or a mix of both. This makes ceramic materials substantially more brittle than metals, which have a metallic bond where valence electrons are shared freely between the atoms of a structure. Therefore, while metals are typically machined in the ductile regime, i.e. a cutting based on plastic material deformation is performed, most commonly the machining of ceramics is done in the brittle regime. Brittle machining causes the material to form vertical cracks when a grinding or lapping grain is engaged and horizontal cracks when the load is removed. The formation of horizontal cracks causes surface chipping, which is the main mode of material removal, but due to the formation of the vertical cracks a substantial amount of subsurface damage is being inflicted. However, if a very small depth of cut is chosen, even very brittle materials may be machined in the ductile regime [5].

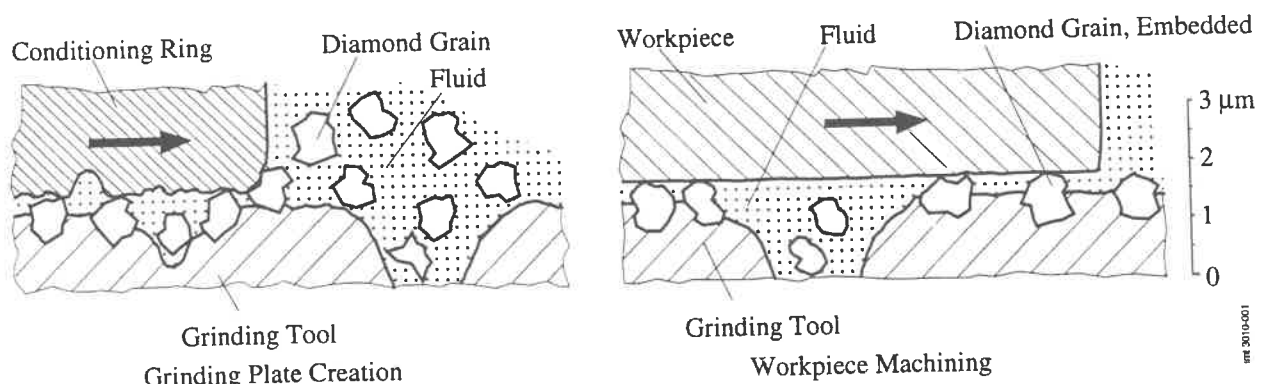


Fig. 1. Nanogrinding process

Nanogrinding is a process based on a lapping kinematic [6]. While lapping uses loose abrasive grain, for nanogrinding the abrasive grain is completely embedded in a soft metallic plate which becomes a grinding tool. Therefore, nanogrinding is a two step process (Fig. 1). The first step is to prepare the grinding plate by embedding the abrasive grain in its surface. In a second step the workpiece is machined. To achieve a supersmooth surface finish, the machining parameters have to be chosen carefully to stay within the ductile regime.

II. GRINDING PLATE CREATION

A. Grinding Plate Conditioning

The grinding plate has a diameter of 400 mm and consists of tin. Early experimental results indicated that nanogrinding cannot be performed successfully on a very smooth grinding plate. Therefore, the plate is conditioned at the beginning of each test cycle using pumice, that consists of alumina (Al_2O_3) and quartz (SiO_2); its grain size is scattering between 1 and 50 μm (Fig. 2a).

Before plate conditioning, the plate is cleaned with DI-water and is subsequently dried by blowing it off with air. To start the 30 minutes long plate conditioning cycle the plate is wetted with a slurry consisting of pumice powder and DI water, mixed at a ratio of 0.25 g/ml. Then the plate is rotated, while a cast iron conditioning ring positioned on the plate also rotates due to friction. The conditioning ring fulfills two functions: (1) it forces the plate's shape to conform to the negative of the envelope dictated by the conditioning ring, thus controlling the plate's macrotopography and (2) it causes the pumice to remove the plate's top layer and to roughen the plate's surface. During the course of the action, the pumice grain breaks down into small pieces, requiring to replenish the pumice slurry at a rate of 30 ml/min. The conditioning ring's hardness has a major influence on the breakdown velocity; by choosing cast iron, which is substantially softer than the pumice components, the pumice breakdown velocity is kept low. Ultimately, as we will find out, a substantial portion of pumice grain does remain in the plate. The amount of pumice embedded on the plate surface and the effect of its presence on the nanogrinding process will be discussed below. After the completion of the pumice plate treatment, the plate is subjected to the cleaning cycle outlined above, thus removing machining debris and loose pumice particles. At the end of the conditioning process the plate's surface roughness R_z is approximately 5 μm .

Plate conditioning has to be done only occasionally to accomplish the desired plate roughness. In our case, a plate conditioning is done at the beginning of each test series to ensure equal starting conditions.

B. Diamond Impregnation

The key process to create a grinding plate is the diamond impregnation. For this step, a slurry consisting of 70% by volume DI-water, 20% ethylen glycol, and 10% ethanol, plus 17 carat/l diamond is distributed evenly over the plate surface. Depending on the application, the grain size is somewhere between 0.5 μm and 3 μm and either monocrystalline or polycrystalline diamond may be applied. Fig. 2b depicts an SEM micrograph of monocrystalline dia-

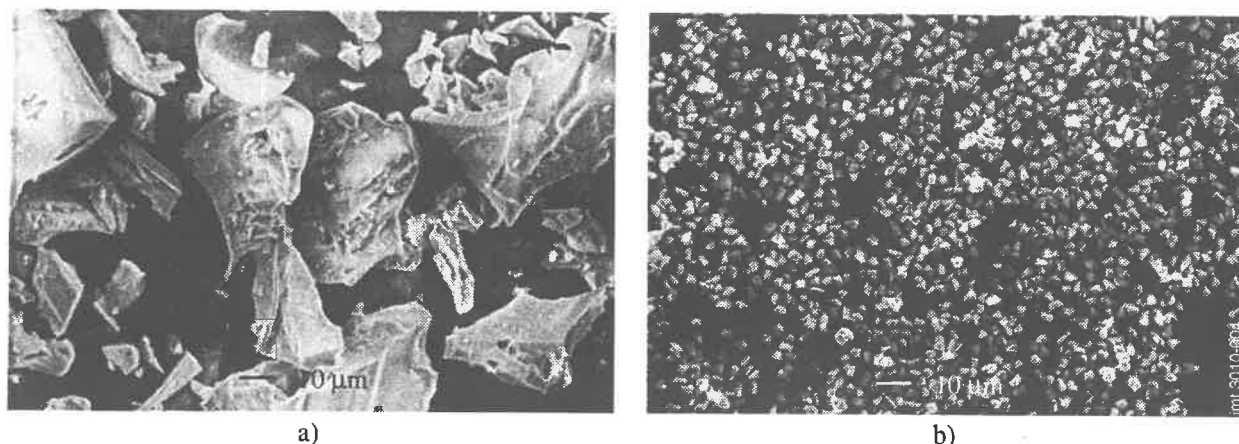


Fig.2. Abrasives: a. Pumice grain for plate conditioning. b. 1.5-3 μm monocrystalline diamond grain.

mond with a grain size of 1.5-3 μm . The conditioning ring utilized for this process step has alumina pellets. The relative velocity between conditioning ring and grinding plate causes the conditioning ring not only to rotate, but also to float on a film of slurry fluid, holding the conditioning ring at a distance from the plate surface. The conditioning ring therefore not only embeds diamond grain in the plate, it also arranges the summits of the diamond grain in a plane coplanar to the conditioning ring envelope. A plate cleaning which removes contaminants and loose, as well as not sufficiently embedded diamond grains, concludes the grinding plate creation.

III. WORKPIECE MACHINING

After grinding plate creation, the actual workpiece machining takes place. As a first step, the plate surface is wetted with the same water based fluid used during diamond impregnation, but for this step the fluid does not contain diamond grain. During machining, fluid is added to the plate at a rate of 1 ml/min to maintain a fluid film of constant thickness. The workpieces are mounted on a disk which is placed inside the conditioning ring. The fluid film allows the workpiece to float on the grinding plate. The fluid film thickness depends on the fluid viscosity, as well as the relative velocity and pressure between plate and workpiece. The optimum working pressure was found to be 0.52 N/mm². It is achieved by adding weights to the disk. This pressure seems to create the proper amount of diamond protrusion to accomplish a machining in the ductile regime. Due to ductile machining, minimal damage is inflicted on the workpiece material's crystal structure at the boundary layer and a superb surface finish can be reached.

Best surface finish results are obtained when applying a two-phase machining process using a staggered diamond grain size. The first phase is a machining with 1.5-3 μm monocrystalline diamond on a grooved grinding plate, followed by a second phase with 0.5-1 μm polycrystalline diamonds on a non-grooved plate. The average surface roughness achieved with two-phase nanogrinding is $R_a=1.14$ nm for Al₂O₃-TiC and $R_a=0.79$ nm for SiC (measured by SFM). Fig. 3 depicts the nanoground SiC surface. The workpiece has a subnanometer average roughness, but the surface finish is slightly compromised by the presence of scratches. Further nanogrinding machining results are presented in [6].

IV. PLATE TOPOGRAPHY AND COMPOSITION

To obtain a good understanding of the nanogrinding process, investigations regarding the grinding plate topography were performed, with the intention to shed some light on the actual machining and plate trueing mechanism. To analyze plate surface topography and composition after conditioning and plate generation, scanning electron microscopy (SEM), combined with energy dispersive spectroscopy (EDS) and scanning force microscopy (SFM) is utilized. Unfortunately, both SEM and EDS are limited to the analysis of a specimen size of below 100 mm which means, that the grinding plate would have to be dissected. For cost reasons, rather than to destroy a grinding plate by cutting it up, the experiment was performed with a tin probe small enough to fit into the SEM. This tin probe was put through a process closely resembling the one for the full size plate.

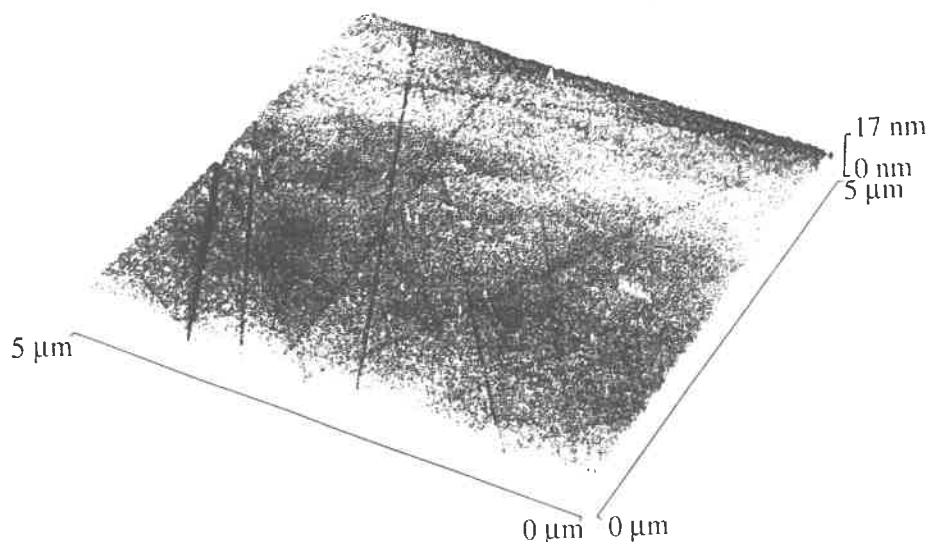
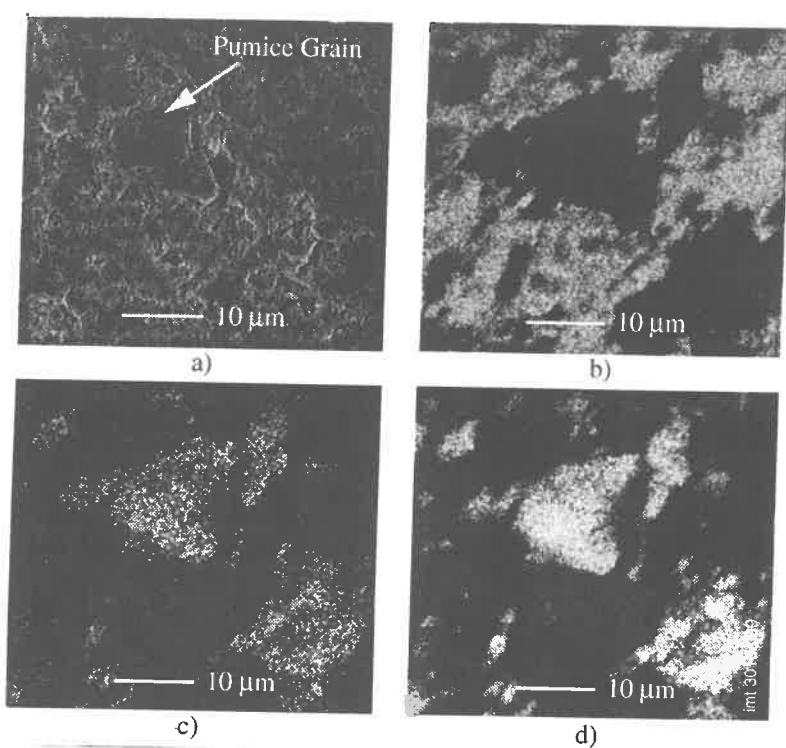
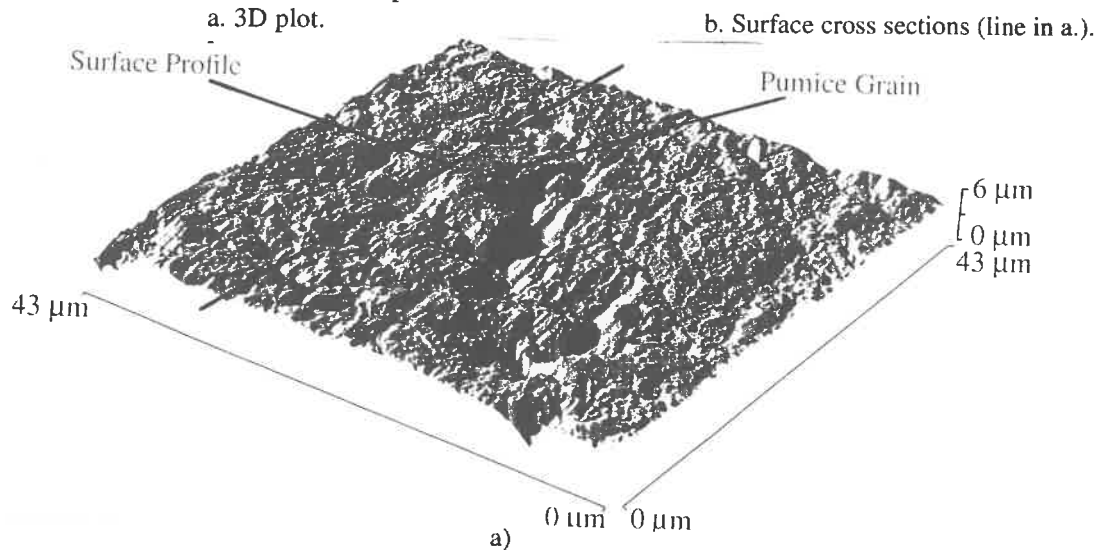


Fig.3. Surface topography of a nanoground SiC probe



Above Fig. 4. SEM/EDS micrograph of a tin probe simulating the grinding plate.
a. SEM micrograph. The arrow points at a pumice particle
b. tin.
c. aluminum.
d. silicon.

Below Fig. 5. SFM scan of the plate surface.
a. 3D plot.



b. Surface cross sections (line in a.).

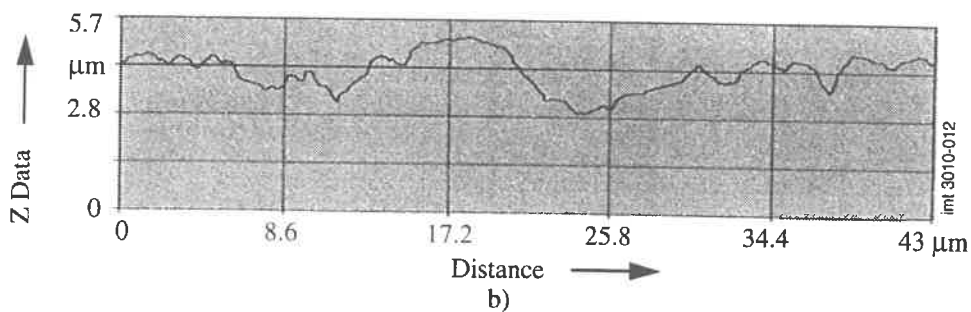


Fig. 4 b, c, and d the EDS scans for tin, aluminum, and silicon. As mentioned above, tin is the plate material, and aluminum and silicon are the elements present in pumice. Bright areas indicate a presence of the respective material, dark areas their absence. The micrographs show very clearly that a major portion of the plate surface consists either of tin or of pumice. Areas displaying the presence of aluminum and silicon are very similar in shape, which is to be expected since pumice, as mentioned above, consists of alumina and quartz. In the SEM micrograph, one of the crystal particles (s. arrow) is clearly noticeable, the respective EDS micrographs characterizes the particle as pumice. All results indicate that a substantial amount of pumice remains embedded in the grinding plate. Fig. 5 depicts an SFM 3D-scan and a surface profile of the vicinity of the pumice grain found previously. The surface profile shows that the presence of the pumice particle substantially contributes to the plate's surface roughness. It also indicates that pumice protrudes above the tin level and that the tin more or less forms plateaus.

While the EDS micrographs provide clear evidence for the presence of pumice at the plate surface, only ambiguous data were obtained regarding the diamond distribution. This may be contributed mainly to the limitations of EDS: it is not capable to analyze elements below atom number 5 (bor) and has only limited capabilities between atom numbers 5 and 11, thus making it difficult to obtain good results on carbon (atom number 6).

Since there is a substantial amount of pumice on the grinding plate surface, the question arises, if pumice plays an active role in the machining process. To answer this question, the machining capabilities of a pumice conditioned plate without diamond impregnation is compared to a plate impregnated with diamond. The diamond used for the second part of this test has a grain size of 1.5-3 μm and is monocrystalline. The workpiece material is SiC, the tests are run over a machining time of one hour each. The removal rate for each test is determined every 10 minutes. Fig. 6 depicts the results, indicating distinctly that the material removal rate of pumice alone is negligible compared to the rate of a diamond charged plate. The decrease in the machining rate over time most dominantly observed in the regular nanogrinding process using a diamond charged plate may either be caused by diamond wear or, more likely, by a settling of the diamond grains in the plate surface due to the vertical machining load.

V. CONCLUSION

To perform nanogrinding successfully, a rough grinding plate is required. The surface roughness is achieved by conditioning the plate with pumice. Pumice particles embedded in the plate surface do contribute to the surface roughness. During the plate generation, the conditioning ring accomplishes three things: (1) it presses the pumice particles embedded during plate conditioning further into the soft metallic plate surface and arranges their summits in

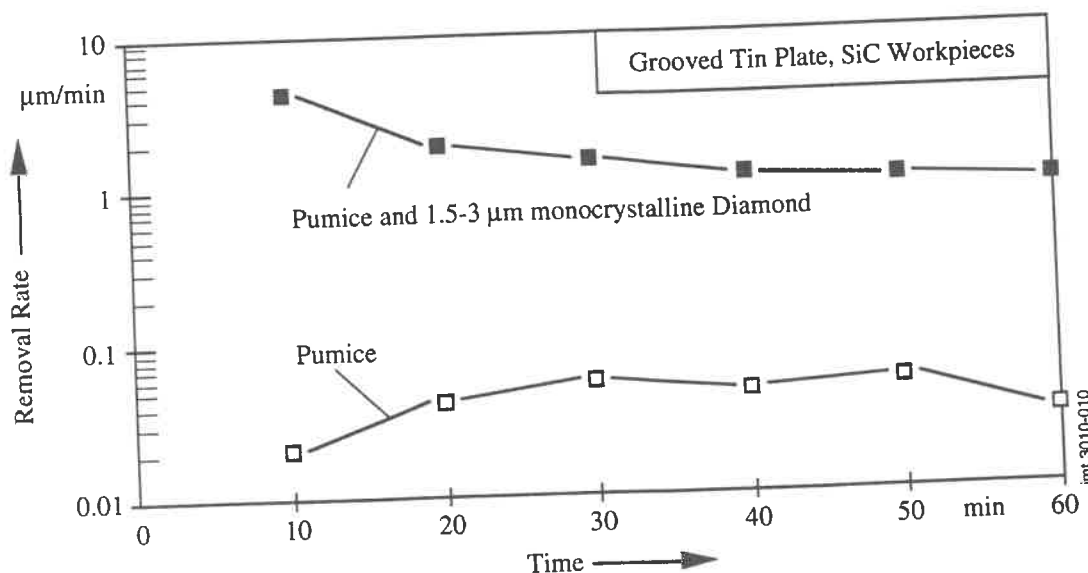


Fig. 6. SiC material removal rate for
a. Pumice.
b. Pumice plus diamond grain.

a plane; (2) it truncates the metal plate's summits in areas where no pumice is embedded, thus creating plateaus slightly below the plane formed by the pumice summits; (3) it embeds diamond grain in the tin plateaus and aligns their summits in a plane. The diamond grain causes a submicron depth of cut at the workpiece and therefore ductile machining. The pumice particles embedded don't contribute in the workpiece cutting worth mentioning, but may help to maintain the plates macroscopic shape required to achieve the required work piece flatness.

Future work will concentrate on a process optimization to eliminate surface scratches as well as investigations how nanogrinding may be applied to the coplanar machining of surfaces on the sandwich structure of materials with various hardnesses.

ACKNOWLEDGEMENTS

This research activity was in part sponsored by the DFG (German Research Foundation) and by Sumitomo Special Metals, Tokyo, Japan, who donated the ceramic materials. Furthermore, the authors would like to express their gratitude to the Hanover Laser Center in Hanover, Germany, for performing the EDS analyses, Topometrix, Santa Clara, California, for providing access to an SFM.

REFERENCES

- [1] H.H. Gatzert, J.C. Maetzig and M.K. Schwabe, "Precision Machining of Rigid Disk Head Sliders", IEEE Trans. Mag. vol 32 (3), pp. 1843-1849, 1996
- [2] H.H. Gatzert and L.P. Ruhbach, "Composite Tripad MIG Head for Proximity Recording", Proc., Idema '96, Santa Clara, California (accepted)
- [3] G. Spur, "Keramikbearbeitung" (Ceramics Machining), Carl Hanser Verlag, Munich, Germany 1989
- [4] C. Chen, S. Sakai, and I. Inasaki, "Lapping of Advanced Ceramics", Materials and Manufacturing Processes, vol. 6(2), pp. 211-226, 1991
- [5] P. Shore, "Damage-free Grinding of Ceramics", British Ceramic Proc. vol. 46, 1990
- [6] H.H. Gatzert and J.C. Maetzig, "Ultrapräzise Planbearbeitung hochfester Keramiken mittels Nanoschleifen (Ultraprecision surface machining of advanced ceramics using nanogrinding)". Schleifen, Honen, Läppen und Polieren, vol. 57, Vulkan-Verlag, Essen 1996 (accepted)

Effect of Vibration Using Ultrasonic and Low Frequency Vibrations on Grooving Process for Hard and Brittle Materials

Hitoshi SUWABE, Ken-ichi ISHIKAWA and Michio UNEDA

Kanazawa Institute of Technology

7-1 Ohgigaoka Nonoichi ISHIKAWA 921 Japan

1. Introduction

In many processing techniques for hard and brittle materials, the mechanical processing using grains is most generally used. Now, the processing methods using grains are classified into two types using free abrasive grains and fixed grains. The method using fixed grains generally uses diamond grains, and it has high machining efficiency as compared with the method using free abrasive grains. However, diamond grains on the tool bring about attritious wear, breaking down and thermal wear, and therefore it has some problems such as deteriorated grinding characteristics.

This study aims to develop a new grooving technique which has high efficiency and accuracy as compared with the conventional technique using a hollow cylinder drill fixed diamond grains (called such as diamond core drill) on the mechanical processing for hard and brittle materials. The new apparatus developed for grooving process is the combined vibration apparatus which can simultaneously supply a diamond tool with ultrasonic vibration and low frequency vibration as shown Figure 1. In the experiment, on the occasion of processing are carried out 4 type methods such as non-vibration, low frequency vibration, ultrasonic vibration and combined vibration (ultrasonic vibration and low frequency vibration) groovings. This paper describes about the effect of these methods as mentioned, the processing characteristics and accuracy.

2. Experimental apparatus and methods

Figure 2 shows outline of experimental apparatus. In this experimental apparatus, the diamond core drill

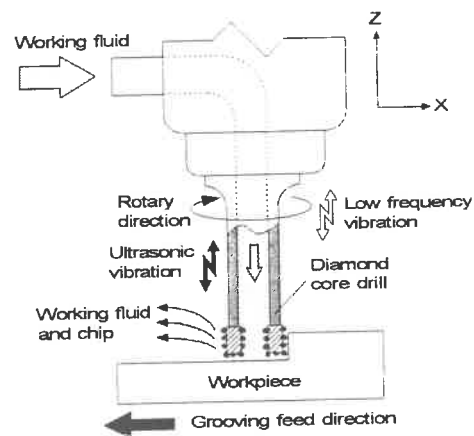


Fig.1 Model of combined vibration grooving

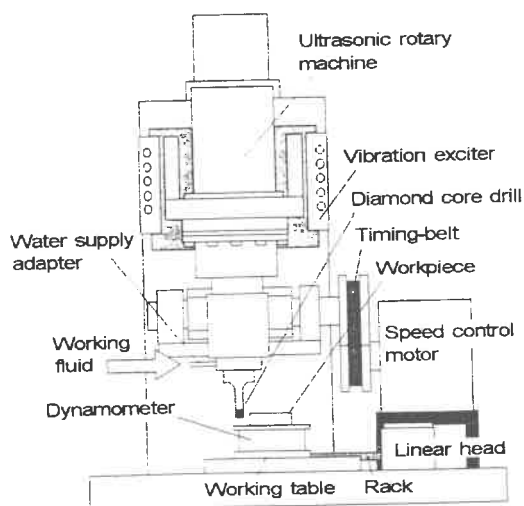


Fig.2 Apparatus of combined vibration processing

is rotated at less than 5000 min^{-1} and is vibrated by ultrasonic vibration (frequency is 20 kHz , amplitudes are $0 \sim 10 \mu\text{m}$) and low frequency vibration (frequencies are $0 \sim 40 \text{ Hz}$, amplitudes are $0 \sim 160 \mu\text{m}$). The diamond core drill is vibrated by each vibrations, which have low frequency vibration and ultrasonic vibration individually, or combined vibrations in the processing. Tap water is used as the working fluid and the fluid is supplied into processing area through the tool inside. The experimental conditions are shown in Table 1.

3. Processing characteristics

Figure 3 shows the relationship between the applied amplitude and horizontal force. The horizontal force of the non-vibration grooving shows by symbol $\triangle$ on Y axis, and another symbols $\triangle$ are results of low frequency vibration. Symbol $\bullet$ on Y axis is a result by grooved of ultrasonic frequency vibration and another symbols $\bullet$ mean combined vibration grooving. As shown in experimental results, the horizontal forces of low frequency vibration and ultrasonic vibration and

Table 1 Main experimental conditions

Diamond core drill	Outer diameter	mm	$\phi 6$
	Inner diameter	mm	$\phi 4$
	Bond, Grain size		Metal-bond, SD#200
Ultrasonic vibration	Frequency f_u	kHz	20
	Amplitude a_u	μm	10
Low frequency vibration	Frequency f	Hz	10 ~ 30
	Amplitude a	μm	15 ~ 470
Number of revolution		$N_0 \text{ min}^{-1}$	2000 ~ 5000
Feed speed of working table		$V \text{ mm/min}$	1 ~ 4
Workpiece	Material, Thickness	mm	Soda glass, 10
Working fluid	Material		Water
	Pressure		29.4

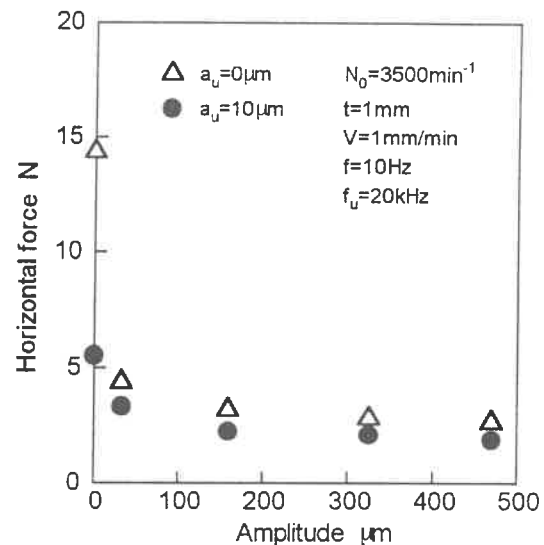


Fig. 3 Relationship between amplitude and horizontal force

ultrasonic vibration grooving are less than one of non-vibration grooving. And, in the case of low frequency and combined vibration grooving, the horizontal force decreases with increase of the applied amplitude at low frequency.

The vibration effect upon grooving is explained by a model of combined vibration in Figure 4. The vibration grooving brings about contact and separate phenomena between the tool and the workpiece in one vibration cycle. Therefore, in the separate time during one vibration cycle, the working fluid can be supplied

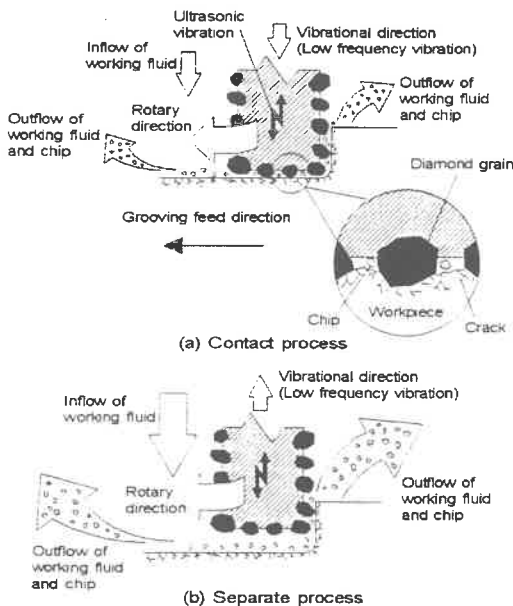


Fig. 4 Removal mechanism of chip in the case of combined vibration grooving

easily into processing parts and removal of chip is able to be improved. These phenomena are clear through measurement of flow rate of working fluid and observations of diamond grains on the tool surface.

Figure 5 shows the behavior of horizontal force obtained by continual grooving process without dressing. In this experiment, the last numbers of each grooving process show the numerical value when the horizontal force becomes 30 N. In the non-vibration grooving, the horizontal force increases rapidly with grooving number and the horizontal force reaches to 30 N at only 10 times of grooving number. The low frequency vibration grooving reaches to the last grooving number at 30 times, and the ultrasonic vibration grooving reaches to it at 70 times. On the other hand, in the combined vibration grooving, the horizontal force reaches to only 20 N over 100 times. Therefore, the tool life is improved in the order of non-vibration grooving and low frequency, ultrasonic, combined vibration groovings.

4. Processing accuracy

Figure 6 shows the relationship between the amplitude of low frequency vibration and

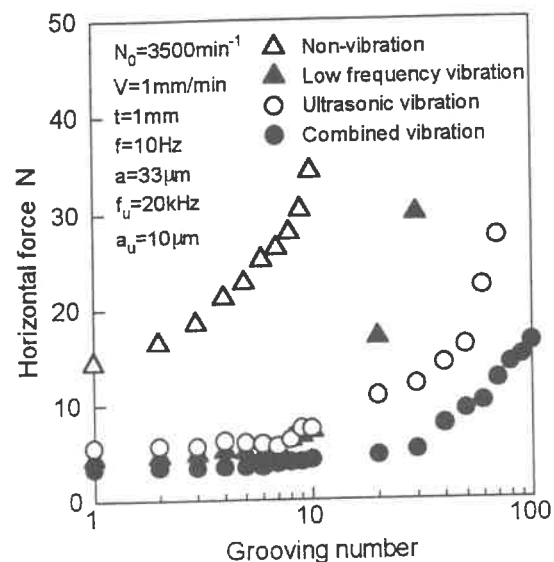
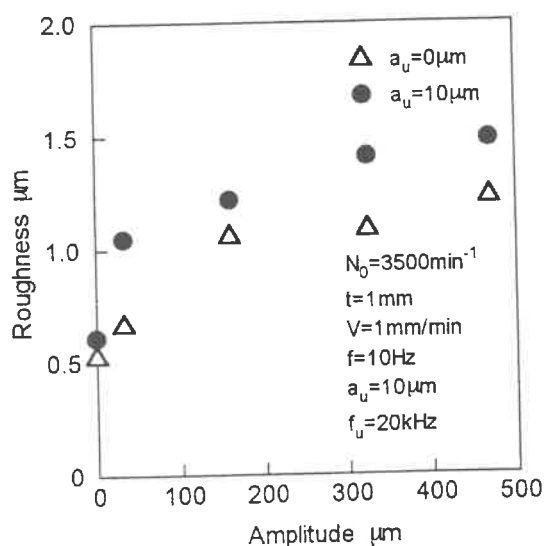
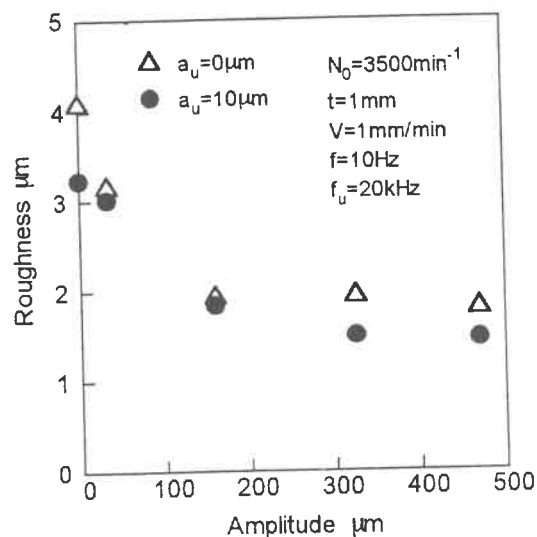


Fig. 5 Relationship between grooving number and horizontal force



(a) Bottom surface of groove



(b) Side surface of groove

Figure 6 Relationship between amplitude and roughness of groove

roughness of grooves. Figure 7 shows microscopic observations of grooved surface. In Figure 7, (a) is the bottom surface of grooves, and (b) is side surface of grooves. The low frequency vibration grooving and the ultrasonic vibration grooving increase roughness of bottom surface. When the amplitude of low frequency vibration increases, the surface roughness becomes still worse. The bottom surface is given continually an impact by diamond grains of the tool. Therefore, this continual impact of diamond grains brings about increase of surface roughness. On the other hand, roughness of the side surface is improved by low frequency vibration, ultrasonic vibration and combined vibration. And, when the amplitude of low frequency vibration increases, the surface roughness is improved. These things are clear from results of Figure 7-(b), especially. That is to say, the side surface of non-vibration grooving and ultrasonic vibration grooving show stripe marks. In the cases of low frequency vibration and combined vibration grooving, stripe marks on the side surface are not brought about.

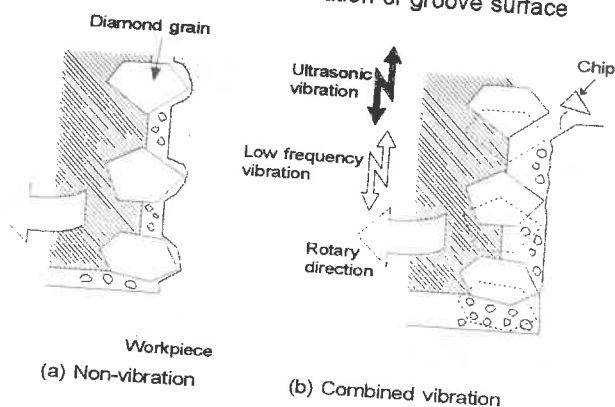
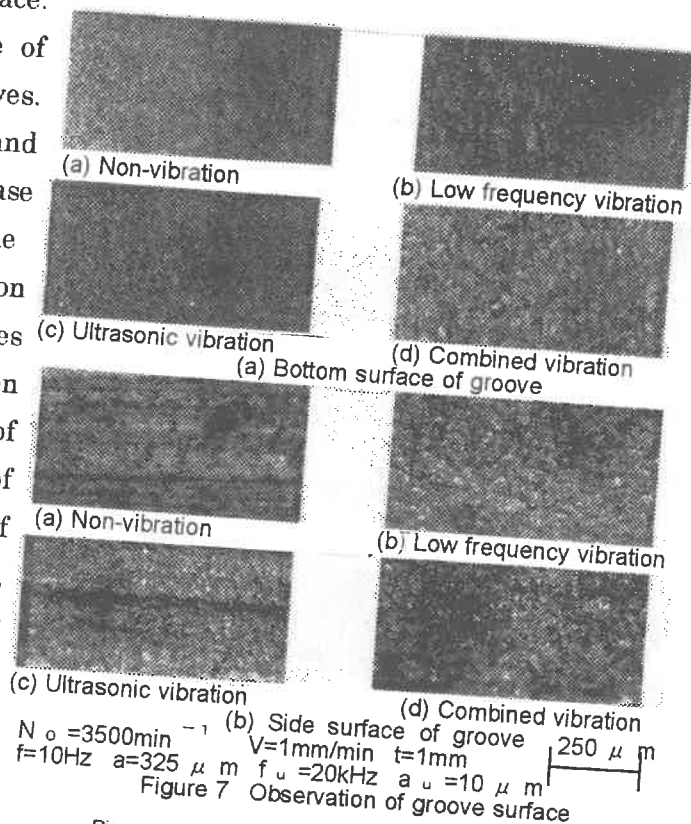


Fig.8 Model of grooving mechanism

Figure 8 shows a model of formation of stripe marks. In the case of very small amplitude, i.e. in non-vibration grooving and the ultrasonic vibration grooving, the stripe marks are formed by diamond grains on the side surface of tool. On the other hand, as the amplitude of low frequency vibration grooving and combined vibration grooving are larger than it of ultrasonic vibration grooving, the stripe marks of side surface are cut.

5. Conclusions

- Through the obtained experimental results, the followings are clear.
- (1) The horizontal force of vibration groovings are smaller than non-vibration grooving.
 - (2) In the combined vibration grooving, the tool life is as much as ten times as compared with non-vibration grooving.

Precision Trimming of Large GFRP Parts Using a Mobile Robot With Onboard Grinder

Yasuo Nasu*, Nobuhiro Kouda**, Kazuhisa Mitobe*,
Toshihiro Kuroda***, Keiichi Shibata*

*: Dept of Mechanical Systems Engineering, Yamagata University
Yonezawa, Yamagata 992, JAPAN

** : Hitachi Yonezawa Electronics Co., Ltd
Yonezawa, Yamagata 992, JAPAN

***: Hitachi Chemical Co., Ltd
Shimodate, Ibaragi 308, JAPAN

1. Introduction

Edge trimming process of large GFRP parts is a remaining process to be automated in GFRP manufacturing because the work space of the trimming task for large GFRP parts exceeds the range of the work space of an articulated robot. We have developed a trimming method using a wheeled mobile robot (WMR) equipped with a handheld grinder⁽¹⁾. In general, control of the motion of WMR is difficult because the existence of a smooth feedback control law, which regulate the motion to an arbitrary point, is not guaranteed⁽²⁾. However, in the edge trimming task, it is not necessary to control the position and orientation of the WMR simultaneously to an arbitrary point. The edge trimming task can be carried out by controlling the WMR to track the desired contour of the GFRP parts.

In this study, a precision curve edge trimming method using a wheeled mobile manipulator equipped with a handheld grinder is investigated. When the robot cuts off along a curve-line trajectory the grinding force due to the removal of GFRP may cause distortion of the grinding disk and slip of the tires. In order to obtain the required accuracy of the products and quality of the surface cut off, control of the grinding force is significantly important. The force control is based on the proper choice of the distance from the center of the disk to the contact point between the desired curve-line trajectory and the side surface of the grinding disk.

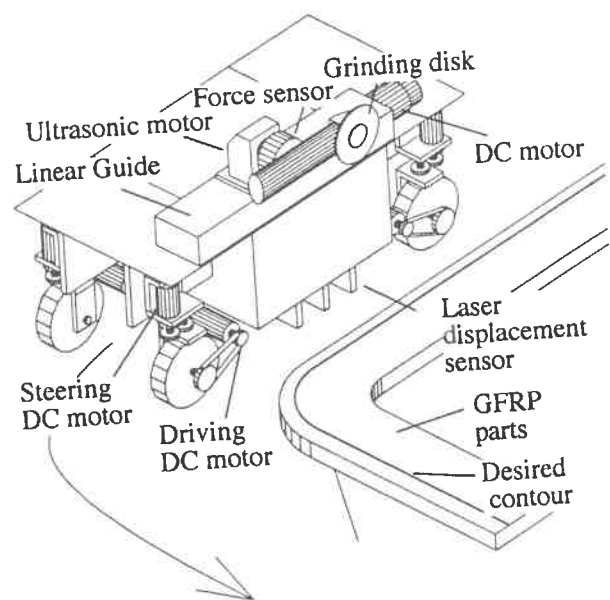


Fig.1 Wheeled mobile robot and GFRP parts.

2. Wheeled Mobile Manipulator System

Workpieces of GFRP, which measure over 2 by 4 m. are easily deformed under the external forces and hence the edge trimming process is necessary to guarantee the required accuracy: $\pm 5\text{mm}$.

In the edge trimming, the wheeled mobile manipulator (WMM) is controlled to track a reference trajectory which is determined by the desired contour of the GFRP parts. The developed WMM system is shown in Fig.1. The WMM consists of the WMR and the 2D manipulator. The WMM is approximately 850mm in length, 500mm in width, and 300mm in height and 57kg in weight. In

order to determine the center of rotation of the WMR arbitrarily on the floor, a four-wheel-steering (4WS) system is adopted in the WMR. The WMR has also a four-wheel-driving (4WD) system and each wheel is steered and driven by two independent DC motors. The 2D manipulator consists of a prismatic joint and a revolute joint. The 2D manipulator is equipped with a handheld grinder. Specifications of the grinder are as follows: rotational speed is 10000rpm, type of the grinding disk is C80M-6, diameter of the grinding disk is 100mm. Vertical height of the grinding disk can be adjusted by the revolute joint. The prismatic joint moves the grinding disk along the direction of y axis of the WMR. The grinding force is measured by a force sensor which is attached between the prismatic joint and the revolute joint. The position and orientation of the WMR is detected by using three laser displacement sensors attached at the side wall of the WMR, that measure the distance from the guide path which is set on the floor. The guide path, which is set up along the desired contour of the GFRP with a curved plate, determines the reference trajectory of the WMR.

3. Distortion of the Grinding Disk

The elastic distortion of the grinding disk due to the grinding force in the lateral direction should be decreased because it causes the inclination of the finished surface.

Fig.2 illustrates the elastic distortion of the grinding disk. Point E is the contact point, where the grinding disk is in contact with the desired contour of the products as the WMR tracks the desired contour. Here V_{ex} , V_{ey} are velocities of the grinding disk and $\dot{\psi}$ is the angular velocity of the grinding disk. In order to reduce the distortion of the grinding disk, V_{ex} is taken to be zero. Since the lateral grinding force varies depending on the position of the contact point on the grinding disk, the inclination of the surface cut off can be reduced by the proper choice of the position of the contact point.

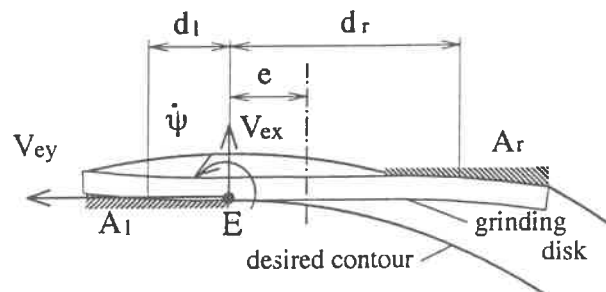


Fig.2 Elastic distortion of grinding disk.

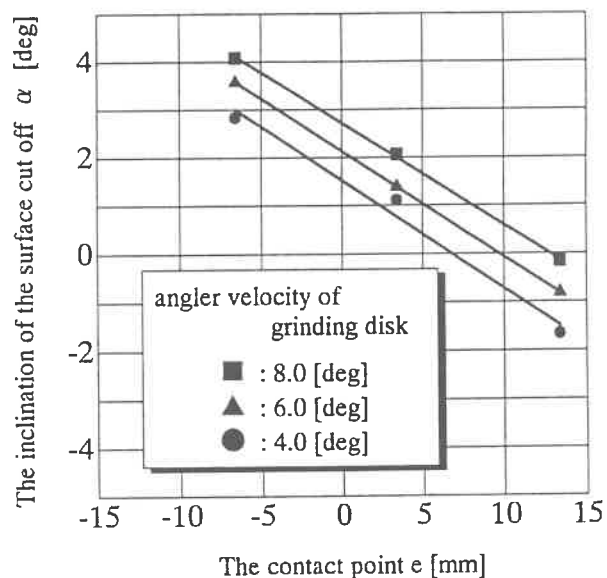
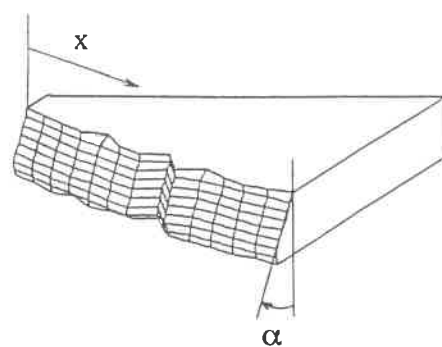


Fig.3 Relation between the inclination of surface cut off and the contact point.

Fig.3 shows the experimental results for the contact point $e = -6.6, 3.4$ and 10.4 mm, where e is the distance of the contact point from the center of the grinding disk. In the experiments, cutting speed of the grinder is 0.04 m/sec, thickness of the GFRP parts is 4

mm and the vertical height of the grinding disk from the surface of the GFRP is 40 mm. The result shows the existence of the optimum value of e and the value varies as the angular velocity of the WMR changes. From the experimental results the contact point is determined accordingly to the next equation so as to reduce the inclination of the surface cut off.

$$e = 0.8 + 1.5\dot{\psi} \quad (1)$$

4. Contour Tracking Control

In this paper, we define a feedback law which locally stabilize the motion of the WMR to track the desired contour. For the simplicity of discussion, we assume that the reference trajectory for the WMR is given by a smooth curve formed by a constant curvature. The coordinate system is shown in Fig.4. ϕ_i and V_i ($i = 1, \dots, 4$) represent the steering angle and the translational velocity of each wheel, respectively. By fixing $V_{ex} = 0$, motion of the WMR is regarded as rotation about a point P , and hence, the velocity and steering angle of each wheel for given θ can be obtained. Defining a new coordinate system Σ_{tri} with its origin at the intersection of the reference trajectory and the line which goes through the contact point E (See Fig.2), the motion of the robot is given as

$$\begin{bmatrix} \dot{x} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} -V_{ey} \sin \psi \\ \dot{\theta} - \frac{V_{ey} \cos \psi}{R+x} \end{bmatrix} \quad (2)$$

Where x is the distance between the contact point and origin of Σ_{tri} , ψ is the orientation of the grinding disk with respect to axis y of the coordinate system Σ_{tri} . R is the curvature of the desired contour. Regarding θ as the control input system(2), a feedback control law is defined by linearizing the system(2) around $\psi = 0$, $x = 0$.

$$K = [K_r, K_\psi] \quad (3)$$

Where $\zeta = [x, \psi]^T$, $K_\psi < 0$, $K_r > -V_{ey} / R^2$.

K_r and K_ψ are determined from simulation

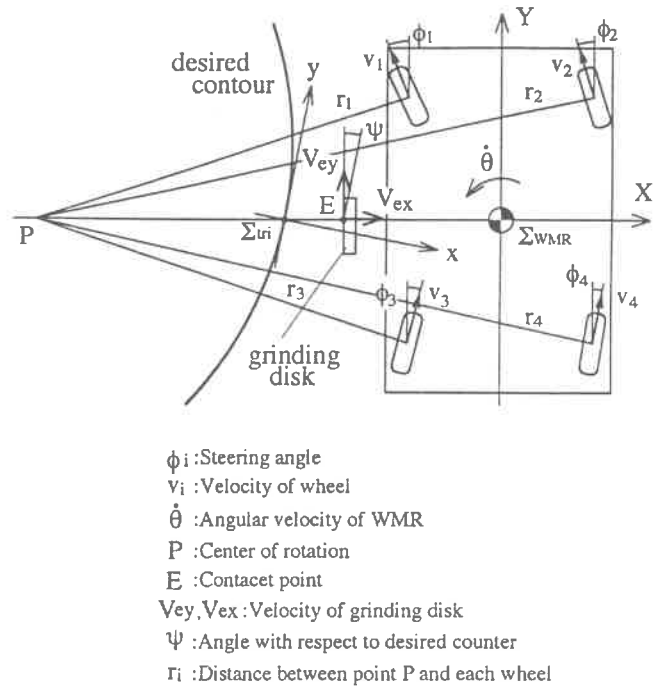


Fig.4 The coordinate systems.

and experimental results.

5. Control Strategy for Trimming

Since the excessive grinding force may cause the slip of the wheels of the WMR, the force should be sufficiently small to maintain the tracking of the WMR to the desired contour of the GFRP parts. In the trimming system, the vertical force can be controlled by adjusting the driving velocity of the WMR because the grinding force is known to be proportional to the removal rate of the material and the removal rate is approximately proportional to the driving velocity of the WMR.

In order to maintain the acceptable tracking for the trimming task, we control the driving velocity of the WMR by using a PI controller, where the vertical grinding force is feed back to obtain the command value for the velocity. Fig.5 shows the block diagram of the trimming robot system. Note that in order to maintain the value of e given by equation (1), the prismatic joint moves the grinding disk along the y direction of the Σ_{WMR} , while the contact point E stays on the

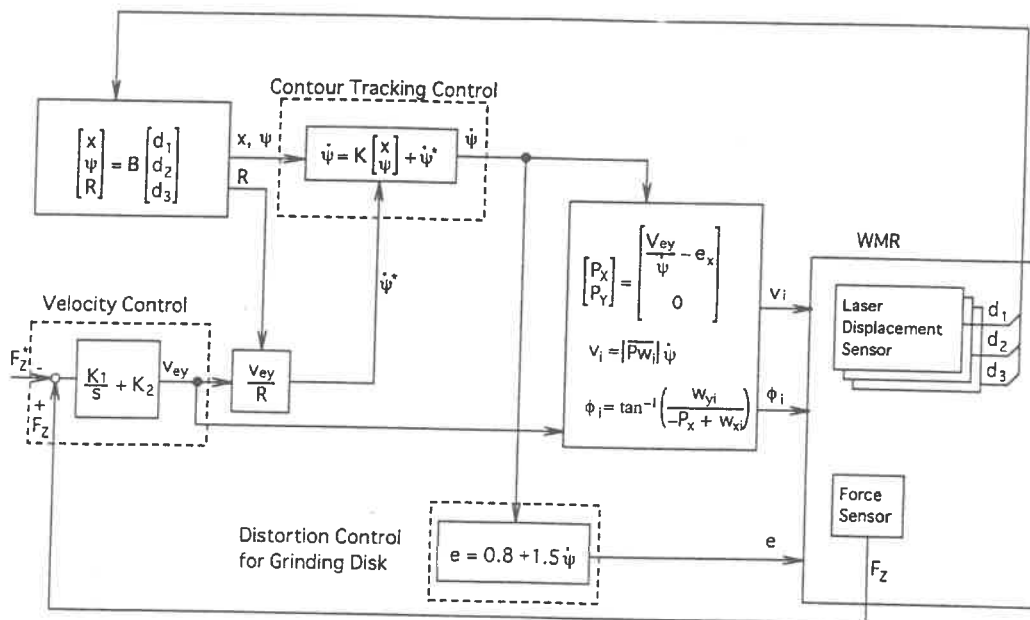


Fig.5 Block diagram of the trimming system.

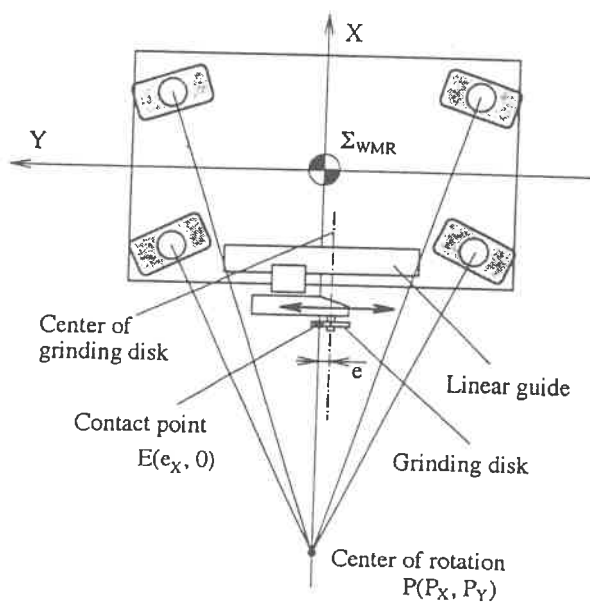


Fig.6 The linear guide to maneuver the contact point of the grinding disk.

straight line connecting the origin of the Σ_{WMR} and the center of the rotation P (See Fig.6).

6. Experimental Results

In order to examine the performance of the trimming system, experimental trimming

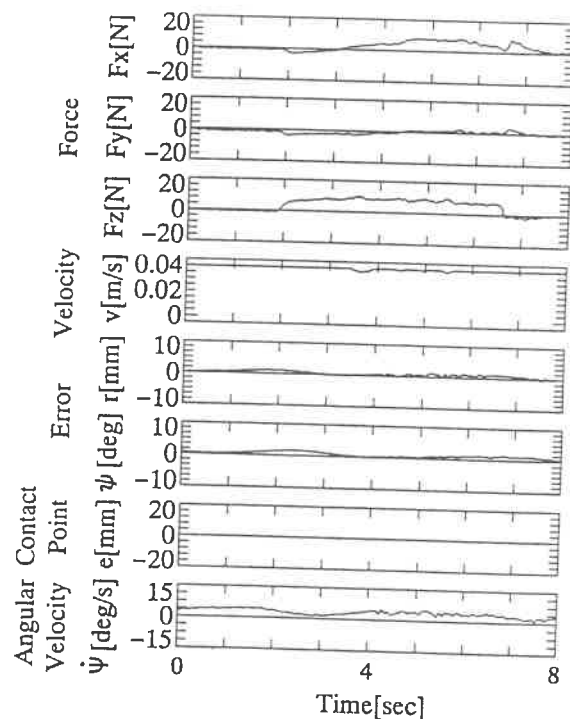
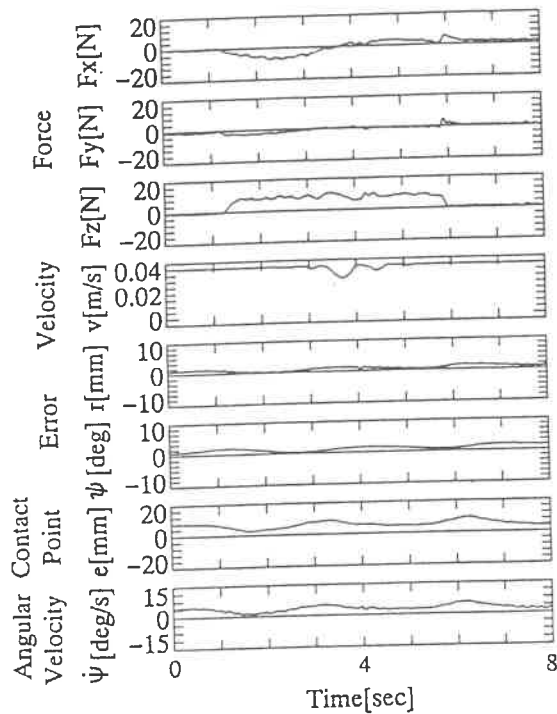


Fig.7 Experimental results without CPCM

tasks were carried out by using a GFRP plate with thickness of 4mm. The radius of the desired contour R_{cut} was chosen to be 500mm



desired contour of GFRP : $R=500.0$ [mm]
 desired grinding force : $F_z=10.0$ [N]
 contact point : control

Fig.8 Experimental results with CPCM

and the guide path with the radius of 560mm. Fig.7 shows the experimental data without using the contact point control method (CPCM). In this case, the tracking error from the desired contour is 1.5mm~3.0mm. The average of the inclination of the surface cut off is $+3 \sim -1$ deg (Fig.9(a)). Fig.8 shows the experimental data with using the CPCM. The contact point distance from the center of the disk varied corresponding to the error of the position r and orientation ψ . The tracking error from the desired contour is 0.8mm~3.3mm. Fig.9(b) shows the average of the inclination of the surface cut off is ± 1 deg.

7. Conclusion

In this study, an automated trimming system by using WMM is developed for the precision edge trimming of large GFRP parts. The inclination of the surface cut off due to the grinding force in the lateral direction of the grinding disk is experimentally analyzed.

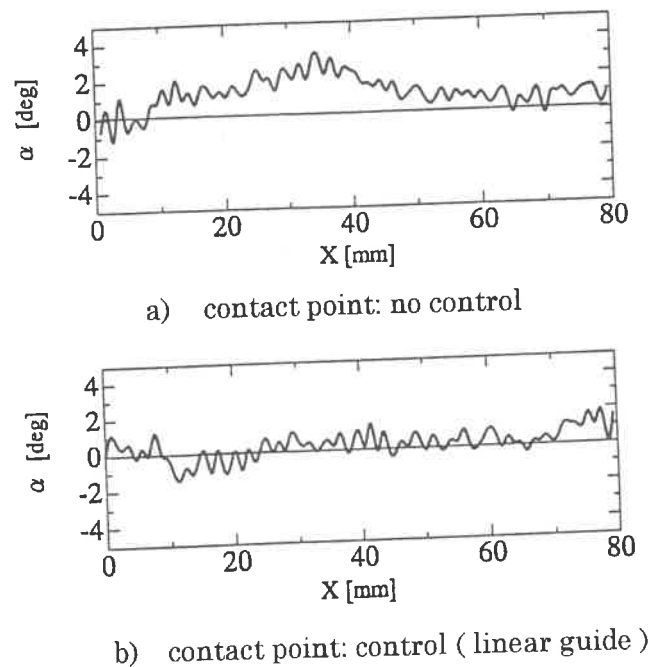


Fig.9 The inclination of the surface cut off.

In order to decrease the distortion of the grinding disk, the contact position between the grinding disk and the desired contour is controlled so as to minimize the lateral grinding force. On our experiments we achieved the required accuracy of ± 5 mm and a small inclination of the surface cut off ± 1 deg. Some experimental results of the trimming task are given to show the effectiveness of the developed system for automated edge trimming of large GFRP parts.

References

- 1) N. Kouda, Y. Nasu, K. Mitobe, K. Shibata: Development of a Trimming System of Large GFRP Parts Using a Mobile Robot, Proc. 2nd IABTEC Conference, 1995, p.489/494.
- 2) C. Canudas and O.J. Sordalen: Exponential Stabilization of Mobile Robots with Nonholonomic Constraints, IEEE Trans. Automat. Contr., 1992, Vol.37, p.1791/1797.

THERMAL CUTTING OF PLATE GLASS

K.Sadashivappa
Research Scholar

M.Singaperumal
Associate Professor

Mechanical Engineering Department
Indian Institute of Technology, Madras - 36
e-mail: mech2@iitm.ernet.in

C.S.Bhaskar Dixit
Lecturer, Mechanical Engineering Department
Bapuji Institute of Engineering and Technology, Davanagere-1

INTRODUCTION

Generally all optical instruments are expected to produce high precision results. Hence, it invites highest precision machining work. The heart of any optical instrument is the lens system. A lens is expected to produce an image which is an exact replica of a given object. The performance of any lens is limited by the presence of defects. Hence utmost care is taken in the fabrication of lenses.

The fabrication of lenses involves different stages which include blank preparation, grinding, polishing and centering and edging. Normally the optical glass is supplied in the form of thick slabs or thinner plates of required thickness. These are cut into squares to the approximate size of the lens required. The corners of the square plates are cut off with a glass cutter. These plates are then cemented on to a flat tool using pitch or a mixture of beeswax and resin. The tool along with the blank is held on a grinder and the grinding is done until the required lens diameter is obtained. An another alternative is to cut glass blanks with a trepanning tool in a standard drilling machine with adequate coolant flow. Then the plate is ground, polished and centered to get a lens of required curvature on the glass surfaces [1,2].

In the above conventional process of blank preparation, the number of operations involved are more and there is a possibility of developing internal flaws and small cracks, which reduce the performance of the lens. In the present paper a new method is proposed for cutting the lenses. The proposed method of thermal cutting of glass plates consists of propagation of a crack in glass, owing to thermal stream produced with the help of a moving hot air jet. A crack, if already present in the glass, can be extended. The movement of the jet governs the direction of propagation of the crack and thus it is possible to cut a glass plate accurately to any desired size and profile in a single step. Hence processes like chipping and grinding to size the blank can be avoided.

THERMAL CUTTING PROCESS

Any scratch or burr on the surface of a glass plate act as part of stress concentration. Hence the glass plate may break or fail easily under the influence of the residual stresses as well as thermal stresses caused by the external changes. This is known as static fatigue of glass which is used to cut glass by creating similar conditions.

Technique of thermal cutting

The hot air jet which is a main feature of this process can be produced by the flow of compressed air through a nozzle. The air is heated up by a heating coil. The crack should be initiated near the edge of the glass plate and can be propagated by moving hot air jet according to the required profile to be cut.

Material selection for the tube

The material of the tube should be such that it sustains the maximum temperature of 325°C required for producing hot air and should not create short circuit problem. Metallic tubes are not preferred owing to their conductivity and so a glass tube shall be of a good choice and being transparent material, glass tube provides a visual control on the performance of the heating element.

Control of air temperature

The voltage and current supply to the heating element governs the air jet temperature. By varying these with the help of a dimmerstat the temperature of air jet can be varied. The approximate value of current ranges between 3 Amps to 5 Amps and the temperature of jet ranges between 150°C to 280°C.

Other features

The compressed air enters at the top of the glass tube and rushes out of nozzle at the bottom in the form of a jet. There should not be any air leakage in order to maintain adequate flow of air. Leakage of air can possibly occur only at three openings (provided for applying voltage to the heating coil and for introducing a thermometer) and this leakage can be minimised with the help of hard rubber stopper coated with a layer of sealing compound. A hole is drilled along the axis of the stopper so that an aluminium electrode can be inserted. A thermometer is inserted to record the temperature of the hot compressed air to avoid over heating of the tube. The operating pressure is to be kept around 0.25 to 0.5 bar.

Parameters

The parameters which influence the thermal cutting process are:

1. Pressure (P), bar
2. Temperature (T), °C
3. Thickness of glass plate (t), mm
4. Type of glass plate
5. Nozzle diameter (d), mm
6. Stand-off distance (x), mm

Pressure and Temperature

The inlet pressure in the heating tube governs the air jet characteristics such as air pressure and air velocity. At higher pressures the force with which the air impinges on the glass plate is more and hence plates of larger thickness can be cut.

The normal temperature of air is around 200°C for obtaining the crack propagation. At higher operating temperatures the thermal stresses induced are more and hence crack propagation is faster. During operation care should be taken such that the heating element does not touch the walls of the glass tube.

Thickness of glass plate and type of glass plate

As the thickness of the glass plate to be cut increases the cutting speed decreases and the operating temperature has to be increased to increase the cutting speed.

Varieties of glasses also do affect this cutting process. Plate glass can be easily cut. Higher temperature is needed to cut tempered glasses.

Nozzle diameter and stand-off distance

A nozzle of diameter about 0.5 mm gives a fine air jet and hence it is possible to have a sharp cut edges. By increasing the nozzle diameter, the jet air will distribute over a larger area and hence the temperature is distributed over a larger area, making it difficult to cut sharp edges.

A very small stand-off distance will cause back pressure in the tube and hence the tube may break. If the stand-off distance is increased beyond a certain value, then the air jet diverges and hence the cutting becomes difficult. Hence an optimum stand-off distance should be maintained.

EXPERIMENTATION, RESULTS AND DISCUSSION

The thermal cutting setup consists of the following: A glass tube (Fig. 1) with a heating element of 1 kW capacity, a dimmerstat to vary the voltage applied to the heating element and hence the air-jet temperature, a mercury thermometer (0-400°C) to measure the air-jet temperature, a x-y table to facilitate the straight line and a rotary table for circular cutting. Nozzles of different diameters are fixed to the tube using a hard rubber stopper. Stand-off distance between the jet and glass plate is varied by moving the glass tube vertically.

Types of cutting

(a) Straight line cutting

When the cut is to be made, an initial crack is obtained about 2 mm in length by a diamond cutter. It should be ensured that air jet is kept in fixed position by properly positioning the glass tube. Then compressed air is injected through the nozzle. A gap of about 2 mm is provided between the glass plate to be cut and the nozzle point. Then voltage is applied to the heating element. Temperature of the air jet can be controlled with the help of the dimmerstat. The hot air jet, first impinges on the area of the initial scratch. After about 10 seconds, the crack begins to propagate in a straight line, when the glass plate is moved in longitudinal direction with respect to the nozzle.

(b) Circular cutting

The glass plate is mounted on a rotary table using suitable fixtures and is fed against the nozzle to get a circular shape.

Experiments are conducted with glass plates of three different thicknesses (3, 4 and 6 mm) and nozzles of two different diameters (1.3 and 2 mm) are used. A pressure of 0.25 bar is maintained. Straight line and circular cuts are made at different temperatures and the time taken for cutting a certain length is noted. Figs. 3 (a-d) show the results of the straight line cutting and Figs. 3 (e-h) show the results of the circular cutting.

Fig. 3 indicates that the cutting time increases with increase in the plate thickness at a particular temperature and stand-off distance. The time decreases with increase in air jet temperature and it is less when the nozzle diameter is 2 mm and also for the stand-off distance of 2 mm.

CONCLUSIONS

A new method for cutting glass plates is reported. Various experiments are conducted on the method and the results are presented. This method can be used for precise cutting of optical lenses and ortho dental photographic mirrors.

REFERENCES

1. Davidson A. (Ed.), Hand book of precision engineering, Vol. 3: Fabrication of non-metals, The Macmillan Press Ltd., London, 1971.
2. Elliott A. and Home Dickson, Laboratory Instruments - their design and applications, Chapman and Hall Ltd., London, 1959.

ACKNOWLEDGEMENTS

The authors are thankful to the following Sriyuths: Balaji Singh, Mallikarjuna Reddy, Siva Punna Rao and Venkateshwara Rao, Ex-scholars of B.I.E.T., Davanagere for their kind help during the experimentation and Karnataka State Council for Science and Technology, Bangalore for the financial assistance to carry out this work.

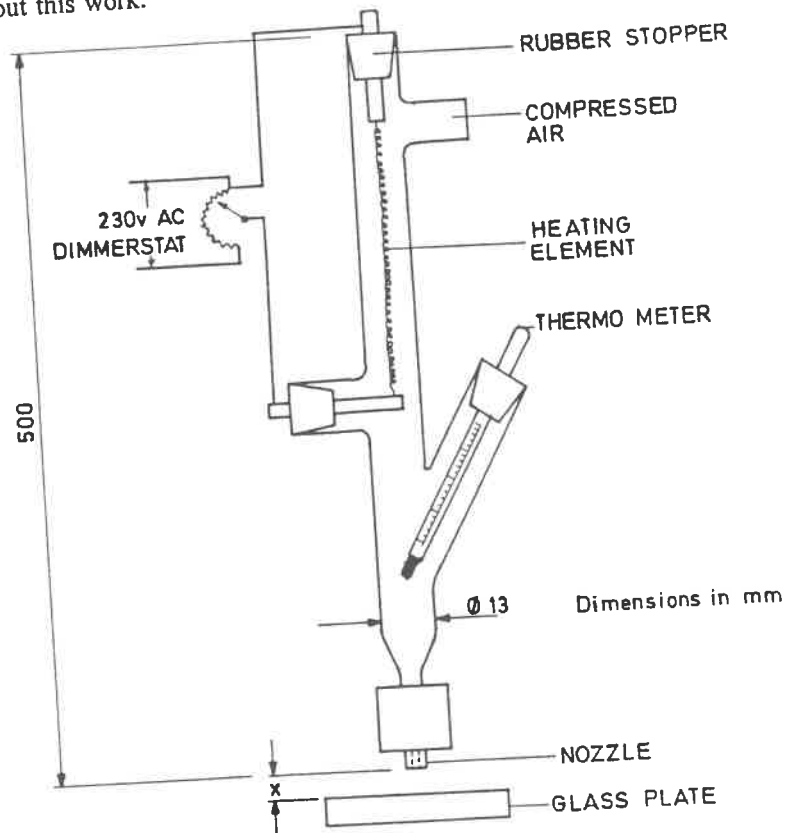


FIG.1 GLASS TUBE

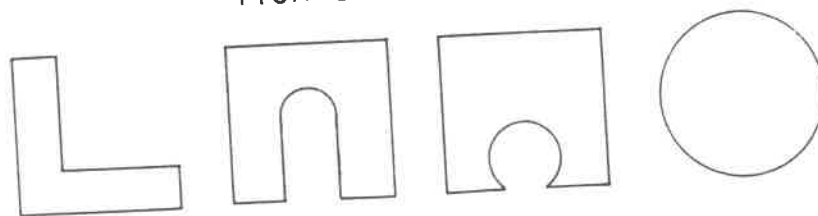
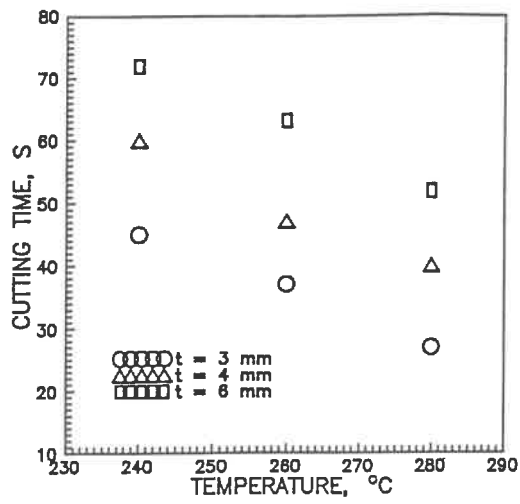
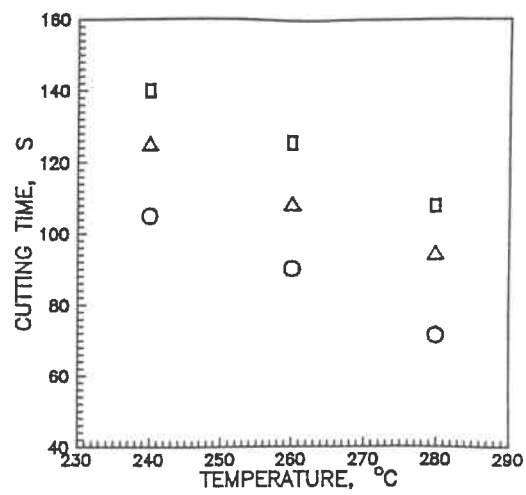


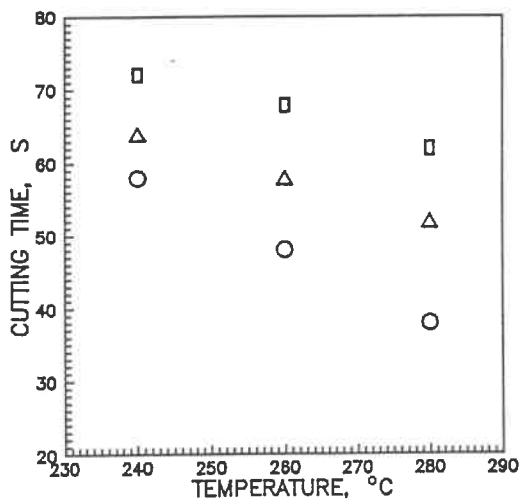
FIG.2 TYPICAL FORMS CUT



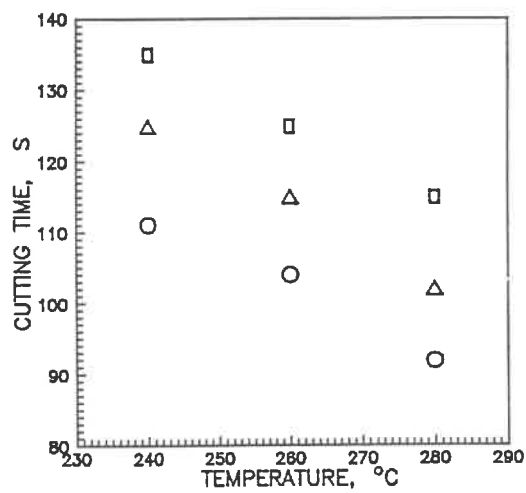
(a) $d = 2 \text{ mm}, x = 2 \text{ mm}, l = 100 \text{ mm}$



(b) $d = 1.3 \text{ mm}, x = 2 \text{ mm}, l = 100 \text{ mm}$

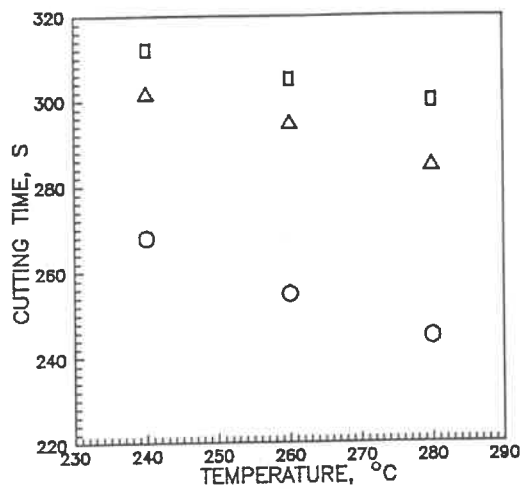


(c) $d = 2 \text{ mm}, x = 1.5 \text{ mm}, l = 100 \text{ mm}$

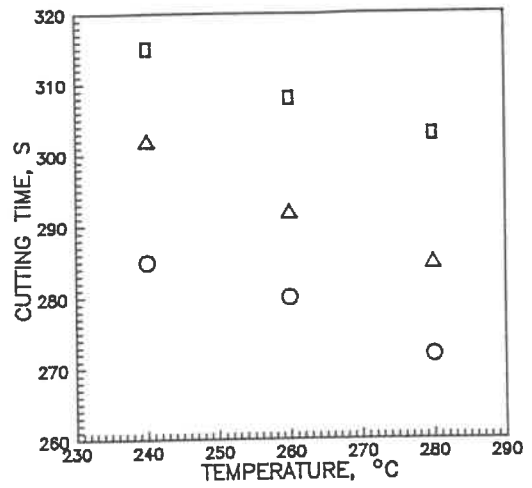


(d) $d = 1.3 \text{ mm}, x = 1.5 \text{ mm}, l = 100 \text{ mm}$

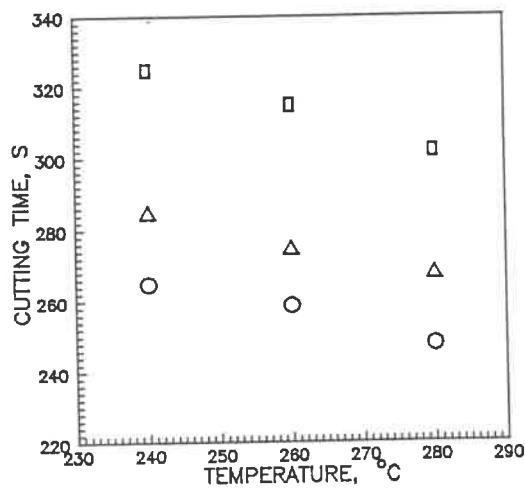
Fig.3 Variation of cutting time with air-jet temperature
 l = length of cut



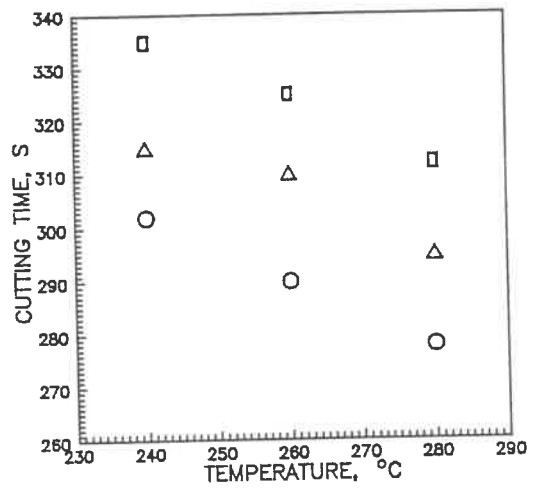
(e) $d = 2$ mm, $x = 2$ mm, dia = 150 mm



(f) $d = 1.3$ mm, $x = 2$ mm, dia = 130 mm



(g) $d = 2$ mm, $x = 1.5$ mm, dia = 130 mm



(h) $d = 1.3$ mm, $x = 1$ mm, dia = 130 mm

Fig.3 (Contd.) Variation of cutting time with air-jet temperature
dia = diameter of circle cut

Precision Polishing of Optical Fiber Connector

Hirofumi Suzuki, Tsunemoto Kuriyagawa and Katsuo Shoji
Faculty of Engineering, Tohoku University
Aramaki-Aoba, Aoba-ku, Sendai, Miyagi, 980-77, Japan

1. Introduction

Recently, advanced optical-fiber signal transmission equipment is being used extensively in various applications. As a result, greater demands are being placed on optical fiber connectors. This paper describes a high-speed precision polishing method for hemispherical surface optical fiber connectors.

The conventional method of machining physical contact (PC) connectors with hemispherical end faces is multi-step lapping using several abrasive lapping films. Use of this method has resulted in some problems in the areas of surface quality, shape accuracy and processing time, in turn causing a significant return.

The proposed machining process consists of three operations, it is composed of three operations. In the first operation, the end surface is ground by a resinoid bonded diamond wheel. In the second, the end surface is polished to hemispherical shape by diamond abrasive. In the final operation, the surface is hybrid-machined using diamond and silica. The new method reduces processing time to short, lowers step height, and reduces the depth of surface damage at the fiber ends.

2. The hemispherical surface optical fiber connector

The hemispherical surface optical fiber connector shown in Fig.1, reduces signal transmission loss but increases return loss in connecting systems. This is due to the fact that Fresnel reflection is eliminated by perfect physical contact (PC) made possible by elastic deformation at the fiber ends. Zirconia ceramic is effective as a connector ferrule because of its high stiffness. Table 1 shows the specifications of the ferrule/fiber end. The radius curvature of the end face should be 10 to 25 mm, the eccentricity of the convex end should be less than 50 μm , and fiber withdrawal from the ferrule end should be less than 0.05 μm through reliability testing under conditions of high temperature and humidity.

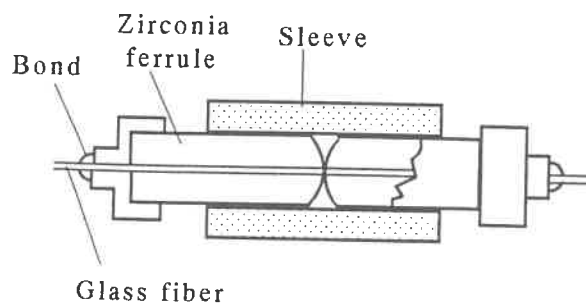


Fig.1 Hemispherical surface optical fiber connector

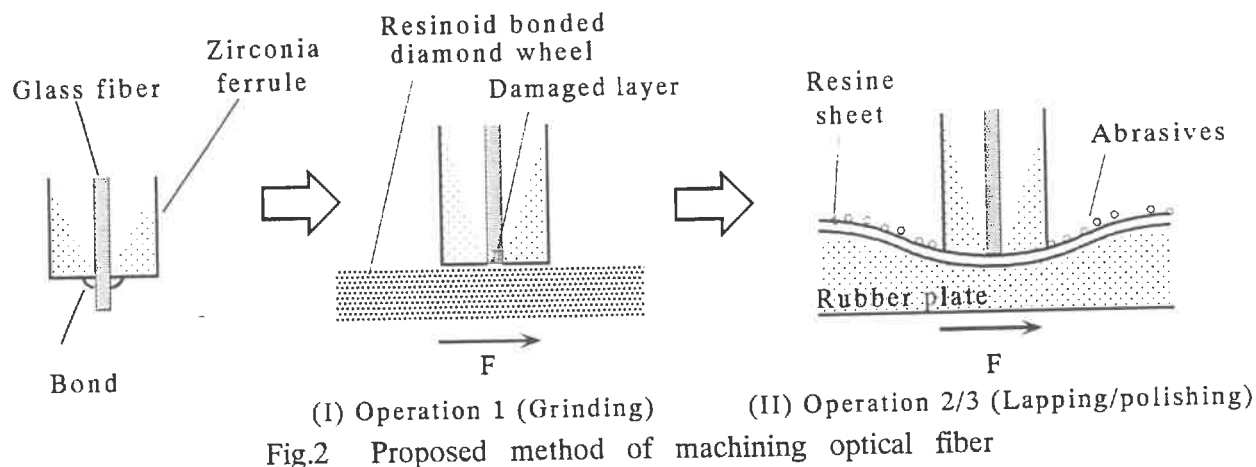
Table 1 Specifications of the PC connector

Radius of ferrule	10 ~ 25 mm
Step height	-0.1 ~ +0.05 μm
Eccentricity	< 50 μm
Insertion loss	< 0.4 dB
Return loss	< -55 dB

3. Machining process

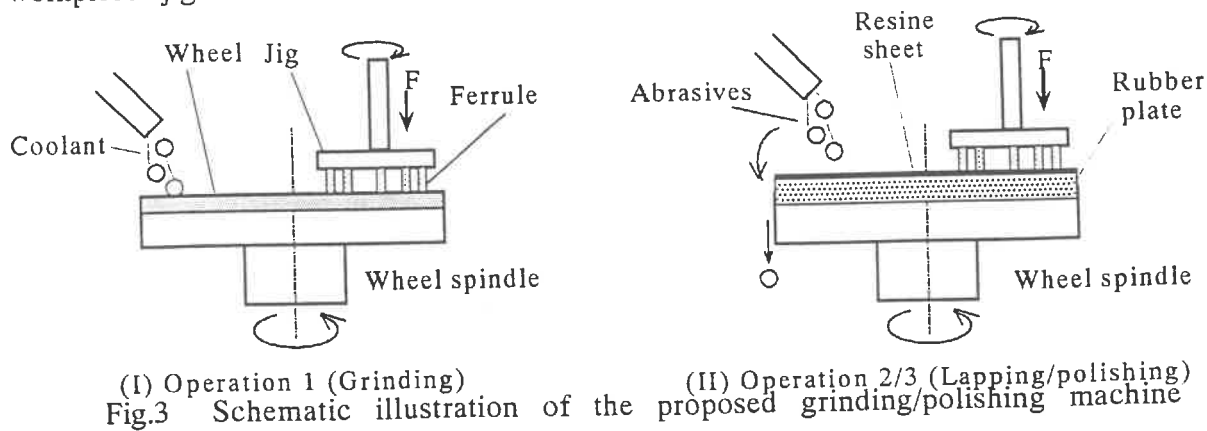
The conventional method of polishing optical fiber requires about 10 machining operations. First, the resinoid bond adhered to the zirconia ferrule end is removed using lapping film. Next, the ferrule end is ground with lapping film on a soft rubber plate to generate a rough hemispherical surface. After that, the end is ground using several grain sizes of diamond lapping film sequentially. Employing this method, processing time is long, and the return loss is significant, a damaged layer remains.

The proposed method requires only three operations shown in Fig.2. In the first operation, the end surfaces are ground flat by resinoid bonded diamond wheel. In the second operation, the end surfaces are lapped into hemispherical shape using diamond abrasives on a layer resine film over a rubber plate. In the final operation, the surfaces are polished with a hybrid abrasives of colloidal silica and diamond on a rubber plate. This hybrid machining is a characteristic operation of the proposed method.



4. Grinding and polishing machine

Fig.3 shows a schematic illustration of the proposed grinding and polishing machine. Several dozens of fibers/ferrules are fixed in holes on a jig plate. Several jigs, weighted in the centers (Fig.4) are set on a grinding/polishing plate. Optical fibers are fixed in each jig. A grinding/polishing disk and the workpiece jig rotate at the same rate and in the same direction.



5. Experimental results

Table 2 shows the grinding/polishing conditions of the three operations. Resinoid bonded diamond wheel was used as the grinding wheel. A 5 mm thick silicon resin plate was used in the 2nd and 3rd operations as the rubber plate. The rotational rate of the grinding/polishing disk was 60 rpm. For the polishing abrasive, 1 μm diamond slurry was used in the 2nd operation, and 80 nm colloidal silica and 1 μm diamond abrasive were used in the 3rd operation.

Fig.5 shows the relationship between the end radius and the polishing load, in the 2nd operation. The horizontal axis is the load applied to one ferrule. As the load becomes higher, the radius becomes smaller. From these results, the load is fixed at 1.90 N.

Fig.6 shows the height of the step between the glass core and the zirconia ferrule resulting from the use of various polishing abrasives in the 3rd operation. In case (a), the polishing abrasive was 1 μm diamond abrasive. In case (b), the polishing abrasive was 80 nm colloidal silica. In case (c), the polishing abrasive was 80 nm colloidal silica and 1 μm diamond abrasive. The step height was lowest in case (a) and within the specified 0.05 μm in case (c).

Table 2 Grinding/polishing conditions

Operation	1.	2	3
Polisher/Wheel	Resinoid bonded diamond wheel	Resine sheel	Resin sheet
Abrasive size	#1200	--	--
Abrasive	Water	Diamond slurry	Colloidal silica
Abrasive size	--	1 μm	80 nm Diamond 1 μm
Load	1.9 N (0.194 kgf)	1.9 N (0.194 kgf)	1.9 N (0.194 kgf)
Rotational rate	60 rpm	60 rpm	60 rpm

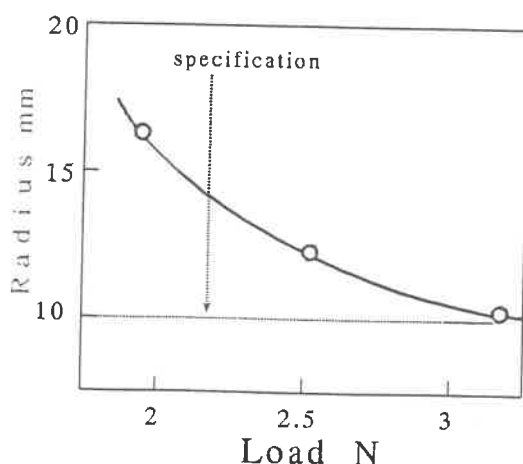


Fig.5 Variation of the end radius with the polishing load

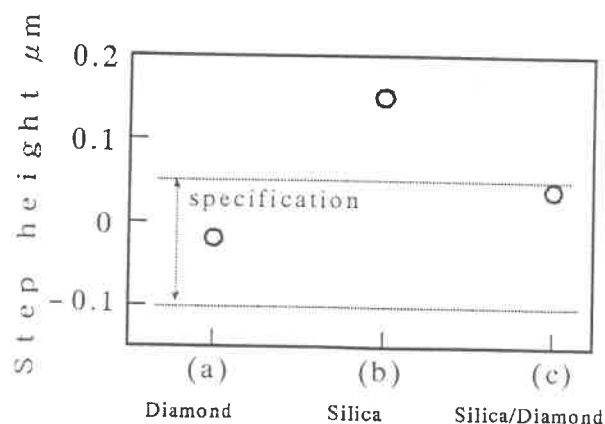


Fig.6 Step height resulting from the use of various abrasives

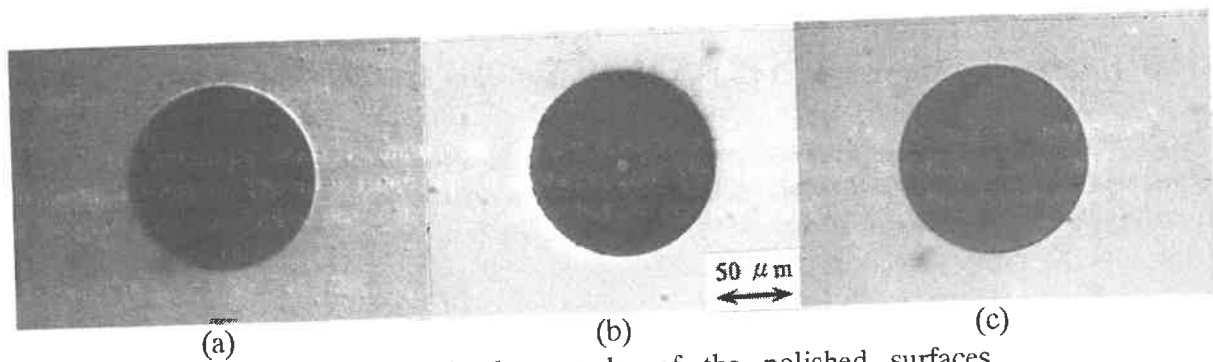


Fig.7 Nomarski photographs of the polished surfaces

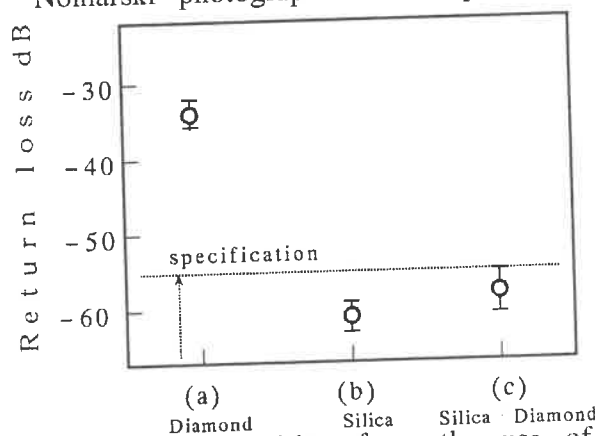


Fig.8 Return losses resulting from the use of abrasives

Fig.7 shows Nomarski photographs of the polished surfaces. Surface (b) is most clear. In (a), some runner cuts are visible. Fig.8 shows the return loss resulting from the use of various abrasives and polishers. In (b), the return loss is low because the layer of damage to the workpiece is thinnest. surface (c), resulting from the use of a hybrid polishing of diamond and silica, is smooth, and its step height is low. It seems that diamond abrasive makes the end shape precise and colloidal silica decreases a damaged layer.

6. Conclusion

In this report, a new method for precision hybrid polishing of optical fiber connectors was proposed. Then an experiment was performed, in which a physical contact-type fiber was polished using the proposed method. From the experiment, the following results were obtained:

- (1) The radius curvature of the fibers was controlled to 15 ± 3 mm, the step height between fiber and ferrule was less than $0.05 \mu\text{m}$, and the eccentricity from the ferrule center axis was less than $30 \mu\text{m}$.
- (2) The insertion loss was less than 0.05 dB and the return loss was to be less than -55 dB.
- (3) The total processing time was less than 15 seconds (per fiber end surface).

References

- 1) J.Watanabe and T.Saitoh, Proceedings of ASPE'89,(1989)133.
- 2) S.Matsui et al., Proceeding of ASPE'95,(1995)179.

In-Process Detection of the Ductile-Brittle Transition in High Speed Scratching of Glass-Theory & Experiment

Jay Daniel, David Dornfeld, Laboratory for Manufacturing Automation, Department of Mechanical Engineering, University of California, Berkeley

Abstract: In the machining of brittle materials, the transition from ductile mode to brittle mode material removal occurs as the material removal mechanism migrates from plasticity dominated mechanisms to mechanisms dominated by material fracture. In process differentiation of the material removal mechanism is an important factor in controlling the productivity and quality of the machining process. Passive acoustic monitoring is shown to yield a robust method of monitoring the ductile brittle transition. Transient acoustic waveforms in the machined material generated by representative ductile and brittle acoustic source functions are discussed. Surface displacements from these acoustic sources are calculated using a Green's function for an infinite plate. Trends in the displacement waveforms which will allow the in-process discrimination of ductile and brittle machining modes in a real machining process are discussed. A high speed transitional ductile-brittle scratching test is performed. Results show that signal peakedness is a strong indicator of the machining regime, per the prediction based on the theoretical waveforms.

INTRODUCTION

The problem of developing an in-process sensor to detect the amount of brittle damage occurring during the machining of brittle materials has interested many investigators. Such a sensor would have applicability in the control of processes involving fine grinding of brittle materials. The sensor most commonly used for these investigations has been the acoustic emission sensor.^{1,2}

Previous work has established empirical relationships between the relative degree of brittle damage seen in the work and the frequency content, specific energy, and the event count rate of the acoustic emission.^{3,1,2} These results, while promising, did not offer a final solution to the problem of developing an in-process sensor, because the relationships established were correlative in nature, and would need to be fitted for each process and setup. A more quantitative approach is warranted.

Fortunately, canonical forms for the transient surface displacements from various types of acoustic source functions exist, and computer programs have been developed to calculate these displacements.^{4,5} In addition, investigators have considered acoustic source functions associated with both ductile and brittle machining.^{6,7} An examination of the displacement wave forms that are generated by characteristic ductile versus brittle acoustic source functions show that fundamental differences in the generated waveforms exist.

TRANSIENT DISPLACEMENT WAVEFORMS FROM CHARACTERISTIC DUCTILE AND BRITTLE ACOUSTIC SOURCE FUNCTIONS

Considering the acoustic sources in grinding from the perspective of an acoustic receiver attached to the workpiece, one can list acoustic sources generated during ductile and brittle grinding that generate acoustic waves in the workpiece. Table 1 specifies these sources.

DUCTILE REGIME GRINDING	BRITTLE REGIME GRINDING
Moments from dislocations	Moments from dislocations
Traction from grinding grits interacting with part surface	Moments from fractures
+ grits translating across surface + stick-slip events between grits and part + grit entrance/exit to/from part + tractions from grit fractures	Traction from grinding grits interacting with part surface
Traction from grinding wheel bond material interacting with part surface.	+ grits translating across surface + stick-slip events between grits and part + grit entrance/exit to/from part + tractions from grit fractures
+ stick-slip events between bond and part + step function tractions from bond fractures	Traction from grinding wheel bond material interacting with part surface.
	+ stick-slip events between bond and part + step function tractions from bond fractures

Table 1: Acoustic Sources in Ductile and Brittle Regime Grinding

Examination of table 1 quickly shows that acoustic sources have been classified either as moments or tractions. This is because the mathematical description of transient surface displacement admits acoustic sources as either moments or tractions. The general expression for the calculation of transient surface motion in a medium resulting from an acoustic source in that medium is

$$u_n(x, t) = \int_0^t M_{kl}(\tau) G_{nk}(\tau - t; x)_{,l} - G_{ni}(\tau - t; x) T_i(\tau) d\tau \quad (1)$$

The normal indicial notation is used with u denoting displacement, M_{kl} denoting the moment acoustic source, G denoting the elastodynamic Green's function, $_{,l}$ denoting the partial spatial derivative, $\frac{\partial}{\partial x_l}$, T_i denoting the traction acoustic source, and τ representing a variable of integration in time. $G_{ni}(\epsilon, \tau - t; x)$ represents the Green's function for displacement in the n th direction at position x and time t in a body resulting from a delta function traction source occurring in the i th direction at time τ and position ϵ in a body. Assumptions in (1) include no body forces acting on the medium and all acoustic sources acting as point sources.

Although lack of space prevents a full treatment of all the acoustic sources listed in table 1, some statements about these sources can be made. First, dislocations in ductile regime grinding of brittle materials have been shown theoretically⁸ and experimentally⁹ to produce acoustic waves too small to detect in a typical machining setup. The other tractions associated with ductile regime machining either have frequency content too low to consider in acoustic emission monitoring, or can be treated as tractions with a step function temporal behavior. That is,

$$T_i(\tau) = T_i u(\tau) \quad (2)$$

where T_i is the i th component of the traction vector on a surface of the machined part imposed by the grinding wheel and $u(t)$ is the step function.

Moment sources from fractures, specifically mode I surface breaking cracks, have been studied experimentally⁷ and shown to be of the form,

$$M_{kl} = D \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \hat{S}(t) \quad (3)$$

when the normal to the crack surface points in the x direction and with D being

$$D \approx 20 A_f G_c \quad (4)$$

where D is the magnitude of the moment tensor, $S(t)$ is the time dependence of the moment source, A_f the final crack area, and G_c the Griffith energy. $S(t)$ can reasonably be assumed to be a step function over the frequency range

of interest.

Having identified the characteristic acoustic source functions for ductile and brittle machining modes, the transient surface displacement waveforms that these acoustic sources generate can be calculated. Figure 1a,b shows these displacements in an infinite plate for a sensor position of (50.8,0,0) mm given a source position of (0,0,0).

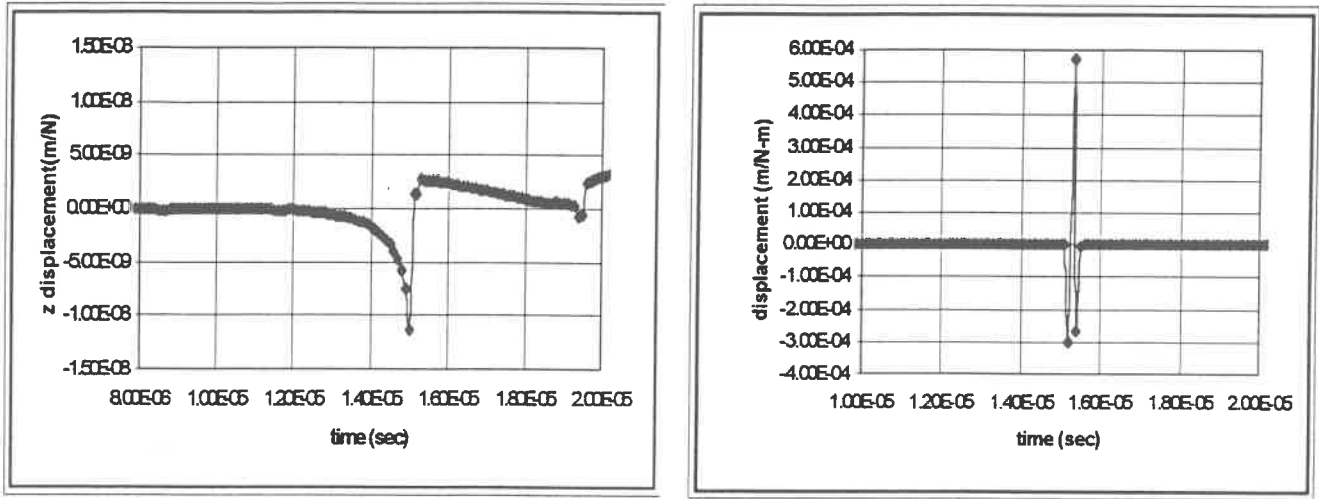


Figure 1a,b: The displacement waveforms in an alumina plate from characteristic ductile and brittle acoustic source functions.

Some characteristics of these signals are notable. First, they occur over very short time periods, requiring the use of a higher frequency acoustic receiver(1-3 Mhz) to monitor them. Monitoring at these higher frequencies also allows the examination of individual wave passages in more detail. For the purpose of discriminating ductile from brittle machining, a signal characteristic from these displacement waveforms is sought that is invariant to the signal energy and is easy to calculate in real time. The peakedness of these signals, relative to their energy, is such a characteristic.

EXPERIMENTAL-HIGH SPEED SCRATCHING

A high speed scratching experiment was devised with this in mind. Figure 2 is a diagram of the experimental setup.

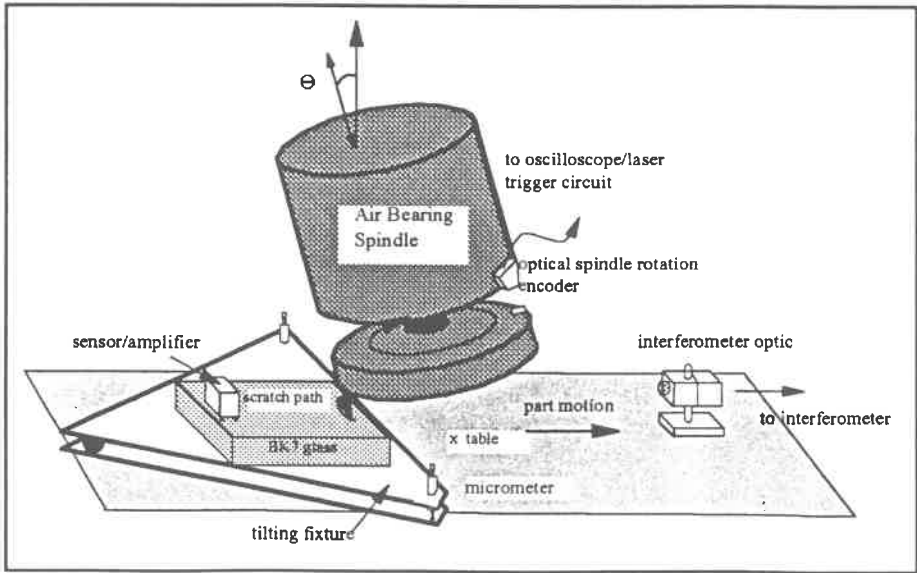


Figure 2: The experimental setup for acquiring acoustic emission from high speed scratching.

In the experiment, a Vicker's indentation tip with a 1x1 um flat on the tip was used to scratch a borosilicate glass (Schott BK7) at increasing depths of cut. Acoustic data was collected with a piezoelectric transducer, the Pinducer™, available from Valpey Fischer (Hopkinton, MA) incorporated into a custom amplifier. Data was sampled at 10 Mhz during the scratching. Each record consisted of 25000 points, representing 2.75 cm of scratch distance at the scratching speed of approximately 9.1 m/sec. The scratch tip position and acoustic data were correlated with a spindle position sensor. The position of the individual scratches on the workpiece and the acoustic file numbers were correlated with axis position data from a linear interferometer.

Scratching was conducted so that the scratch depth transitioned from ductile to brittle over a series of scratches according to the critical depth of cut parameter¹⁰. The scratch numbers were correlated with the acoustic files gathered. Axis position data and tip position on the sample showed that the acoustic file number was related to the scratch number by

$$y = 0.4 * \text{file_number} - 22.84 \quad (5)$$

with all positions in mm. Optical examination of the scratched workpiece with a Leitz optical microscope and a Zygo interferometer showed that the first scratch for which acoustic data was generated was scratch 75 (corresponding to acoustic file #75). This scratch was approximately 5 nm deep. Scratches 76 and 77, along with scratch 75, were ductile in appearance. Scratch 78 was transitionally ductile-brittle. Scratches 79 and on showed largely brittle characteristics.

Because the acoustic sensor had been previously determined to be a surface velocity transducer, and the displacement record was of most interest, the raw signal (which is proportional to velocity) was numerically integrated. Also because previous examination of the acoustic waveform from indentation results showed that most of the significant signal characteristic of crack sources was at frequencies above 1.0 Mhz, the acoustic data files were processed through a high pass filter (9th order Chebyshev) at 1.0 Mhz.

Then, a simple algorithm to characterize the peakedness of the scratch records was developed. The algorithm functioned as follows. For a given scratch record, a 15 sample (1.5 μsec of data) data window was interrogated starting with t=0. The difference between the highest and lowest values in that window was stored. The window was then incremented 0.1 μsec forward, and the maximum difference for that window was found and compared to the maximum difference from the first window. The largest value was retained. The window was incremented through the entire file, looking for the greatest dipole. This value (fig. 3b) was called the maximum dipole of the given scratch record. The integrated/high pass filtered voltage records (proportional to surface displacement) were interrogated in this way.

Next the RMS energy (fig. 3a) of the acoustic signal from the integrated, high pass filtered data files were calculated according to (6).

$$\text{RMS} = \left(\frac{1}{T} \sum_i (u_i^2 \Delta t) \right)^{\frac{1}{2}} \quad (6)$$

where T is the total sample period, Δt is the time increment between samples, and u_i is the volt-sec value proportional to surface displacement. While a smaller window for the total T of the window could be chosen, for initial simplicity, each 2.5 millisecond record was examined separately.

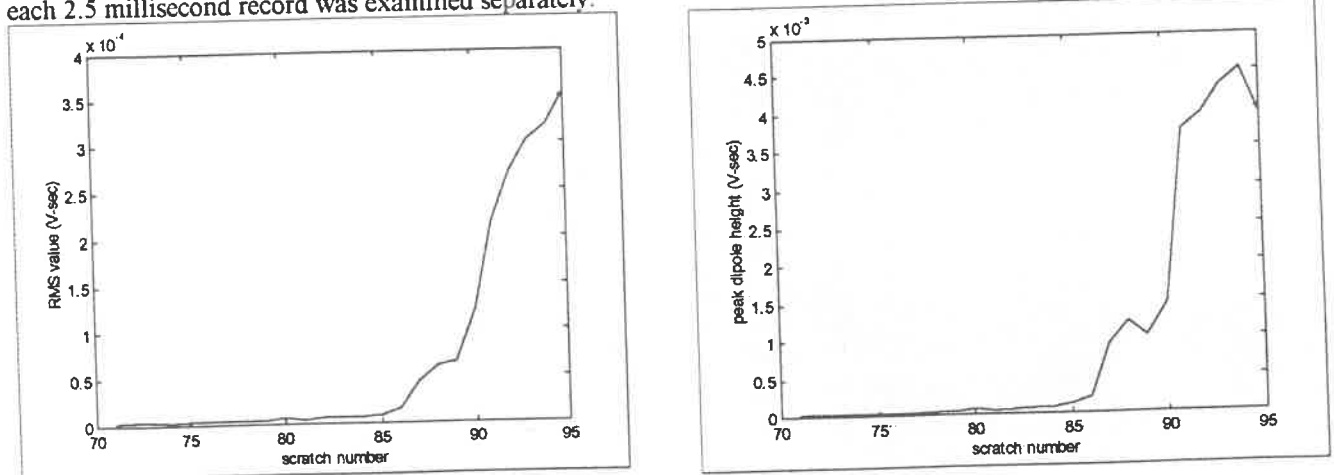


Figure 3a,b: RMS voltage and peak dipole values from the integrated, high pass filtered acoustic records. Records 71-74 are included to allow benchmarking vs. signal noise.

Finally, since it was desired to normalize the peak dipole value to the signal energy, the ratio of the peak dipole values to the rms voltage was taken (fig. 4).

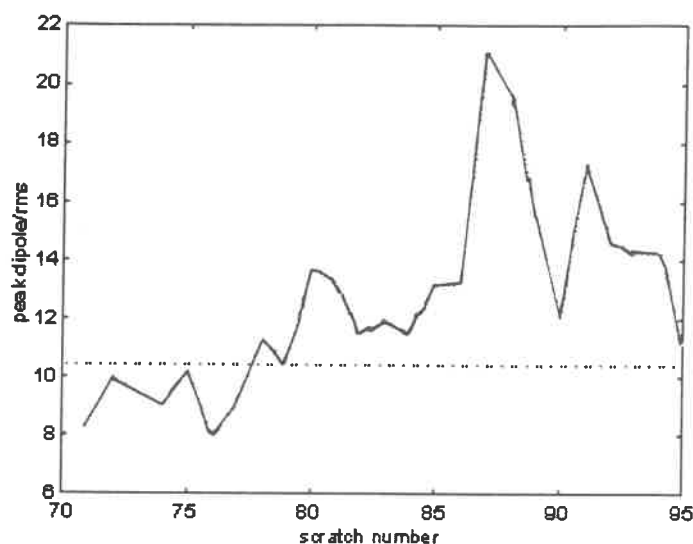


Figure 4: The ratio of peak dipole to rms voltage for scratches 71-95. Again, 71-74 indicate signal noise. Below the line scratches were ductile in appearance. Above the line, they appeared brittle.

This measure of signal peakedness showed a clear transition from ductile to brittle machining correlating exactly with physical observation of the scratches and model of critical depth of cut¹⁰ for the ductile brittle transition. Furthermore, the peakedness of the records 71 through 77 was dominated by noise spikes from the motor driving the x axis. The ratios for records 71 through 77 would have been in the range of 5-7 had this noise source been excluded. The ratio for records 78-95 had no contribution to peakedness from motor noise. It is concluded, then, that an even stronger transition from ductile to brittle machining would have been seen with some modifications to the acoustic receiving system electronics to reduce noise.

CONCLUSIONS

A methodology for detection of the ductile-brittle transition in precision grinding has been developed from consideration of the surface displacement waveforms generated from characteristic ductile and brittle acoustic sources. A high speed transitional ductile brittle scratching experiment was conducted which confirmed that signal peakedness is a robust measure of the ductile brittle transition in high speed scratching. The methodology is invariant to the signal energy, and is motivated by the physics of acoustic propagation in the system.

REFERENCES

- ¹ Akbari, J., Saito, Y., Hanaoka, T., Enomoto, S., JSME International, Series C, **38**(1), 1995, pp. 175-180.
- ² Bifano, T.G., Yi, Y., **14**(4), Oct. 1992, pp. 219-228.
- ³ Ichida, Y., Kishi, K., Hasuda, Y., JSPE, **58**(5), 1992, pp. 841-846.
- ⁴ Johnson, L.R., Geophys. J. R. Soc. **37**, 1974, pp. 99-131.
- ⁵ Hsu, N.N., NBSIR 85-3234. Available through NTIS, PB86-143956/AS (1985).
- ⁶ Hsu, N.N., Simmons, J.A., Hardy, S.C., Mater. Eval., **35**, 1977, pp. 100-106.
- ⁷ Kim, K.P. Sachse, W., J. Appl. Phys. **59**(8), April 1986, pp. 2711-2715.
- ⁸ Daniel, J.P. Ph.D. Thesis, University of California, Berkeley, 1996.
- ⁹ Daniel, J.P. Engineering Systems Research Center Report No. 30, November 1994, Berkeley, CA.
- ¹⁰ Bifano, T.G., Dow, T.A., Scattergood, R.O., ASME Trans., J. of Eng. for Ind., **113**, May 1991, pp. 184-189.

On-Line Measurement of Grinding Wheel Wear Using Acoustic Emission

W.P. Dong

L. Annecchino

J.A. Webster

Center for Grinding Research and Development
University of Connecticut
Storrs, CT 06269-5119
USA

Abstract

A new method based on an application of acoustic emission for on-line measurement of grinding wheel wear is proposed in this paper. The feasibility of the proposed method and the uncertainty of wheel wear measurement are studied with experiments.

1. Introduction

As grinding remains one of the most important precision manufacturing processes, feedback control for high efficiency, high quality and low cost production is of great interest to manufacturing industries^[1-6]. In order to achieve reliable and cost effective feedback control, a critical issue is to realize on-line measurement of grinding wheel diameter which is affected by wheel wear, wheel expansion and dressing operations. Diameter measurement of a grinding wheel can help to quantify wheel wear, to determine feedback compensation in dimensional control, and to decide dressing conditions. So far, however, there are no effective commercial methods available for on-line measurement of wheel wear or wheel diameter. Coolant and swarf in the grinding process forms a hostile environment for advanced measuring techniques, e.g. optical sensors^[5,7]. The cost, the amount of retrofitting work to install a measuring system, the efficiency and the accuracy of such measurement are also big concerns before the system to be accepted by industry. Recently, Oliveira et al^[8] has proposed a convenient method for dimensional measurement of grinding wheel using acoustic emission (AE). However, since a single point dresser was used in their method, the measuring accuracy is highly influenced by local geometry of the wheel, e.g. waviness, roughness and local cracks. In addition, low level AE caused by turbulence and elastic contact in the initial dresser/workpiece contact was very much the concern of their study, which would induce unstable AE detection. Bearing the above constraints in mind, a simple on-line wheel dimension or wheel wear measuring method is proposed in this paper. This method is based on utilizing the CNC controller and acoustic emission measurement. The principle of wheel wear measurement using acoustic emission and experimental approaches are introduced in this paper. The feasibility of the proposed method is experimentally verified and the uncertainty of wheel wear measurement is evaluated. It is expected that the proposed method will be cost effective, reliable, efficient and potentially applicable to manufacturing industry.

2. Description of Method

The proposed method is based on detection of an acoustic emission signal. As it is known, a grinding process produces high frequency (above 100KHz) AE waves due to plastic and elastic deformation of part and wheel materials. It has also been shown that AE is more sensitive than force, acceleration and power to reveal initial wheel/part contact in the grinding process^[6]. Even if an initial contact is so weak that only a few active wheel grits rub against the part, the generated AE signal is still detectable with

sufficient signal to noise ratio for a high level of confidence^[6]. Therefore, the AE measuring technique can serve as a touch trigger probe to be applied for wheel diameter measurement. The core principle of this method requires a reference surface located at an appropriate position on the grinding machine. Whenever dimensional measurement of the wheel is required, the wheel is touched on the reference surface and the corresponding AE signal produced triggers the positional capability of the CNC grinding machine, which compares the current value with the previous reading, to obtain the diameter change in the wheel surface position. This process is similar in operation to a coordinate measuring machine (CMM). The function of a mechanical touch trigger probe of a CMM is replaced by the AE sensor and the function of the CMM controller is replaced by the CNC controller. The accuracy of wheel wear measurement using this method will be influenced by several factors which will be discussed later.

3. Experiments

In order to verify the feasibility and accuracy of the proposed method, experiments were conducted on a CNC surface grinder. The experimental set-up is shown in Fig. 1. Since the controller's positional capability was not accessible during this experimental stage, a displacement sensor - LVDT instead of CNC controller was used to measure the wheel position. For obtaining quantitative wheel wear and its relationship with the measured results, wheel wear was simulated by a truing process, i.e. the amount of dimensional change of the wheel taken by a truing process is considered as the amount of wheel wear. The computer takes readings of two channels, AE and downward displacement of the wheel (LVDT), concurrently. When the wheel touches the reference surface and the root-mean-square (RMS) signal of AE exceeds a preset threshold, the computer records the instantaneous position of the wheel, which is compared with the data obtained in the last measurement and serves as the reference for next measurement. During each measurement, the reference surface was contacted (lightly ground) within a small local area, therefore the table should be moved a short distance horizontally to keep a fresh contact area for touching for next measurement. The experiments used an alumina oxide wheel, with an initial diameter of 420 mm and a width of 11 mm. The horizontal dimension of the reference surface was 100x20 mm, which was initially ground flat in-situ prior to the test.

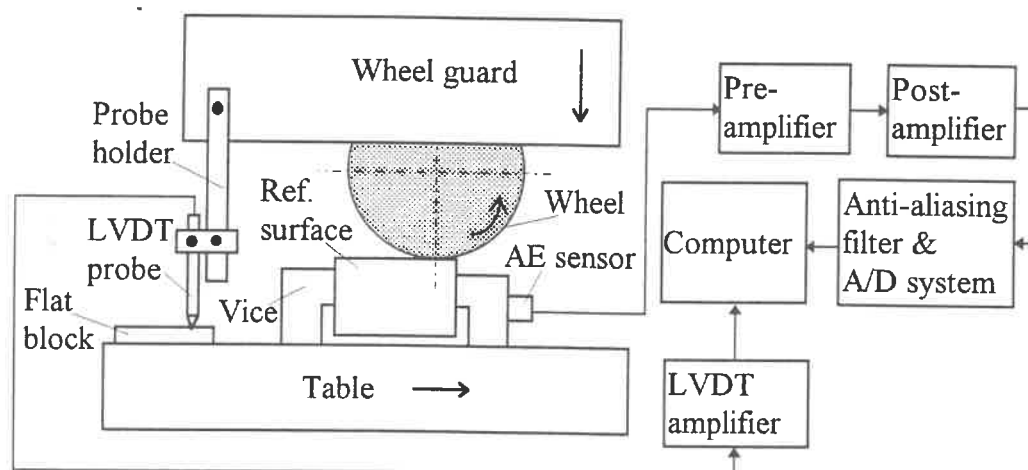


Fig. 1 Experimental setup of wheel measurement using acoustic emission.

With the described setup, two experiments were conducted. One was a repeatability test where the wheel touched the reference surface several times without grinding (except for the time of touching), therefore inducing no wheel wear. This test can help to determine measuring consistency since no wheel wear has occurred. The second experiment was a wheel wear test using the truing process to change the diameter of the wheel, instead of extensive grinding tests to generate natural wheel wear. Fig. 2 shows an example of AE RMS and LVDT signals where the horizontal broken lines are the threshold of AE and the circles represent the moment of AE triggering and corresponding displacement reading. The time constant of the AE RMS signal used was 12.8 ms. As can be seen in Fig. 2a, a larger threshold (20 dB) was set in the experiment. This is to avoid unstable touch triggering caused by a few protrude grits in initial wheel/part

contact and to maintain consistent touch triggering due to substantial wheel/part contact. The wheel or LVDT displacement shown in Fig. 2b has three distinctive zones after the LVDT touches the flat block, i.e. (i) the wheel approaches the reference surface with a fast speed where a linear LVDT displacement is observed (200-240 ms in Fig. 2b); (ii) the wheel approaches the reference surface with a reduced speed where a non-linear LVDT displacement is observed (240-600 ms in Fig. 2b); and (iii) the wheel stops where no further LVDT displacement is observed (flat response, after 600 ms in Fig. 2b). In the third zone, the differences of LVDT readings in consecutive wheel wear tests can reflect position accuracy (or truing compensation accuracy) of the wheel after truing.

4. Analysis of Experimental Results

The experimental results are presented in two phases, i.e. repeatability test and wheel wear test.

4.1 Repeatability of measurement

The repeatability test was conducted by touching the reference surface several times with the wheel rotating but without grinding (except for the moment of touching). After each measurement, the table was moved a few millimeters horizontally in order to offer a fresh reference area for the next measurement. As is seen from Figs. 2b, the wheel approached the reference surface fast and decelerated when it was close to the reference surface, where the wheel displacement is nonlinear due to the servo control. At the time when the displacement reading was taken, the plunge rate was approximately $300 \mu\text{m/s}$. In Fig. 2b, readings were taken from two places. One was obtained from the circled place, which was AE triggering point, another was obtained from the flat response area (after 600 ms in Fig. 2b) where the wheel stopped downfeed and the reading represents the position of the wheel after truing, which should be directly related to the truing compensation value used.

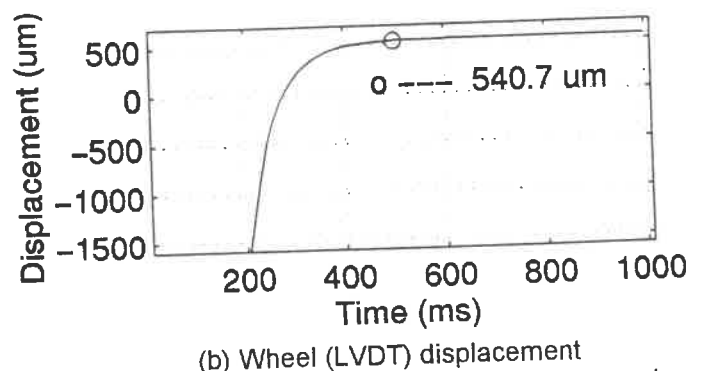
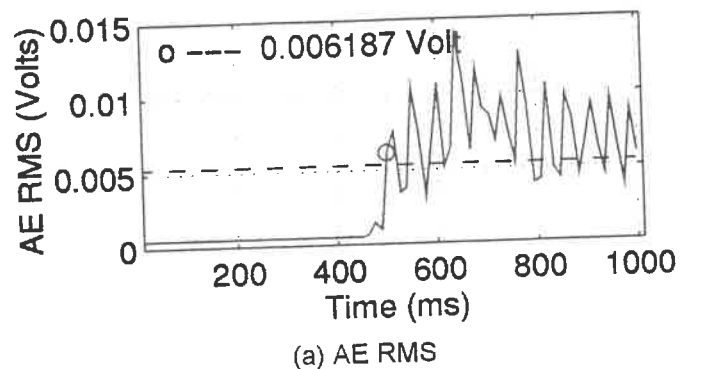


Fig. 2 An example of AE RMS and displacement recorded concurrently.

The data obtained from seven consecutive measurements is presented in Table 1. The deviations of these consecutive measurements are shown in Fig. 3. The uncertainty of the repetitive tests triggered by AE is $U_{r1}=1.9 \mu\text{m}$; the uncertainty of the wheel repositioning is $U_{r2}=2.07 \mu\text{m}$. It appears that the proposed method was as accurate as wheel repositioning at a wheel approach rate of $300 \mu\text{m/s}$.

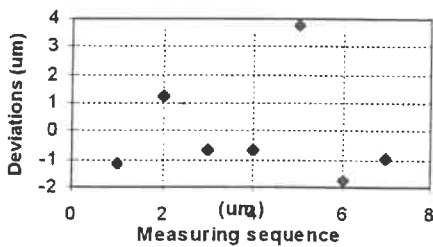
4.2 Wheel wear test

As stated earlier, during the wheel wear test, wheel wear was simulated by the truing process. Between each pair of measurements, the wheel was trued by $25.4 \mu\text{m}$ (0.001 inch). The experimental results for the wheel wear test are given in Table 2 and Fig. 4. Fig. 4a shows the dimensional change of the wheel radius detected by the proposed technique versus the nominal accumulated truing depth; Fig. 4b is the deviations between them and Fig. 4c is the measured step increment after each truing cycle. The step increments are close to the nominal value which is $25.4 \mu\text{m}$. The difference between the measured value

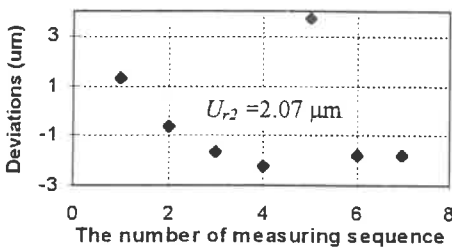
and the nominal value can be attributed to (i) error of the method; (ii) error of wheel reposition and (iii) error between the nominal truing depth and actual truing depth. The deviation of wheel reposition with respect to the nominal value after truing is shown in Fig. 4d. The uncertainty of wheel reposition, U_{w2} , is $1.29 \mu\text{m}$. Surprisingly, the uncertainty of the measured accumulated wheel wear, U_{w1} , is $0.81 \mu\text{m}$, which is smaller than U_{w2} . However, note that the deviations between the measured wheel wear and the accumulated truing depth (Fig. 4b) has a positive bias, which implies that the measured wheel wear is slightly larger than the nominal accumulated truing depth.

Table 1. Data of repeatability test (in μm)

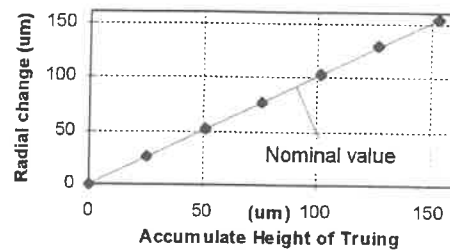
	Measured value	Deviation with mean	Reposition	Deviation with mean
1	539.9	-1.13	551.6	1.82
2	542.3	1.27	549.7	-0.08
3	540.4	-0.63	548.6	-1.18
4	540.4	-0.63	548.1	-1.68
5	544.8	3.77	553.6	3.82
6	539.3	-1.73	548.4	-1.38
7	540.1	-0.93	548.5	-1.28
Mean	541.03	0	549.78	0
STD	1.9	1.9	2.07	2.07



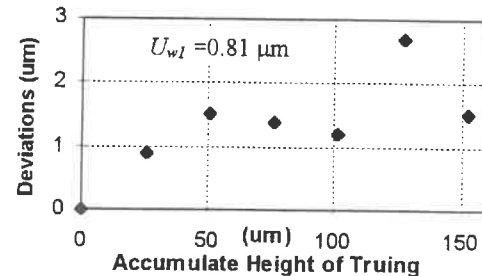
(a) Deviations obtained by AE triggering



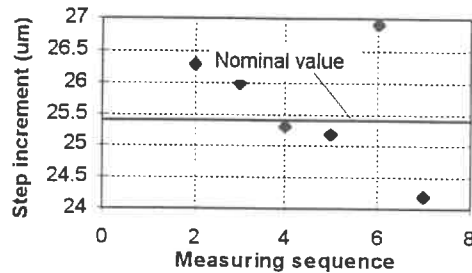
(b) Deviations of wheel reposition
Fig. 3 Deviations of repetitive measurements.



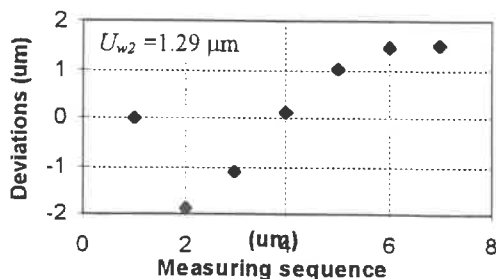
(a) Radial change of the wheel after truing



(b) Deviations between measured radial change and nominal accumulated truing depth



(c) Measured step increment after each truing cycle



(d) Deviation of wheel reposition after truing
Fig. 4 Measurement of wheel wear.

Table 2. Data of wheel wear test (in μm)

	Nominal value	Measured value	Deviation with nominal value	Measured step increment	Reposition value	Deviation with nominal value	Step increment of reposition
1	0	0	0	N/A	0	0	N/A
2	25.4	26.3	0.9	26.3	23.54	-1.86	23.54
3	50.8	52.3	1.5	26	49.71	-1.09	26.17
4	76.2	77.6	1.4	25.3	76.32	0.12	26.61
5	101.6	102.8	1.2	25.2	102.63	1.03	26.31
6	127	129.7	2.7	26.9	128.47	1.47	25.84
7	152.4	153.9	1.5	24.2	153.92	1.52	25.45
Mean			1.31	25.65		0.17	25.653
STD			0.81	0.95		1.29	1.11

5. Discussions of the Accuracy of Wheel Wear Measurement

The accuracy of wheel wear measurement using the proposed method depends on the following factors: (i) the positional accuracy of the grinding machine; (ii) the flatness of the reference surface; (iii) the method to set threshold for AE; (iv) the rate of wheel approach to the reference surface and (v) the response characteristics of the CNC controller. Of these the method to set threshold for AE and the rate of wheel approach to the reference surface are considered the most important.

The application of AE for wheel wear measurement is different from contact detection for gap elimination since in the latter the threshold should be set as low as possible in order to be able to detect early contact of highest grits on the wheel. However, as the wheel consists of waviness, roughness and some protruding grits, the wheel wear measurement would be distorted if a reading was taken when only top or loose grits are in contact with the part surface. Ideally, the diameter reading should represent when the majority of grits are in contact with the surface. To achieve this the DC component of the AE RMS can be used. An example of a significant increase in this DC component is shown in Fig. 2a. Therefore, the threshold of diameter measurement should be set much higher than that for gap elimination. In this report, a threshold of 20 dB was used. Providing this threshold is consistent for two successive measurements, the error will be minimum.

The rate of wheel approach to the reference surface is also an important factor which influences accuracy. For example, assume the wheel approach rate is $1000 \mu\text{m/s}$ and the maximum variation of touching triggering is 10 ms, then a possible measurement error is $10 \mu\text{m}$. Therefore, in order to reduce this kind of error source, the approach rate must not be too large. As from the experimental results, an uncertainty smaller than $2.1 \mu\text{m}$ was achieved with an approach rate of about $300 \mu\text{m/s}$. Therefore it is assumed that a plunge rate smaller than $300 \mu\text{m/s}$ (i.e. $100 \mu\text{m/s}$) would be suitable for practical applications. The slower approach rate need only be applied in close proximity to the reference surface, thus saving a considerable amount of measuring time.

6. Conclusions

Based on the experiments and above discussions, it may be concluded that

- the proposed method can be implemented in practice;
- the experimental results show that under the approach rate of about $300 \mu\text{m/s}$, an uncertainty of smaller than $2.1 \mu\text{m}$ can be achieved in a large range of accumulated diameter change;

- the accuracy of diameter measurement using the proposed method can be as good as wheel reposition;
- the accuracy of diameter measurement using AE can further be improved by properly choosing approach rate and the method for setting the threshold.

In addition an algorithm of the CNC controller should be developed which uses the response of the AE triggering to access the encoder position at each moment of contact, therefore eliminating the need for the LVDT.

References

1. Blum, T. and Dornfeld, D.A., Grinding Process Feedback Using Acoustic Emission, Proceedings of 4th International Grinding Conference, 1990, MR90-525.
2. Konig, W. and Meyen, H.P., AE in Grinding and Dressing: Accuracy and Process Reliability, Proceedings of 4th International Grinding Conference, 1990, MR90-526.
3. Lezanski, P. and Rafalowicz, J., An intelligent monitoring system for cylindrical grinding, Annals of CIRP, Vol. 42/1/1993.
4. Kluft, W., Monitoring the Grinding and Dressing Operations Increases Output and Quality and Reduces Cost and Waste, Proceedings of 5th International Grinding Conference, 1993, MR93-371.
5. Tonshoff, H.K., Werner, F., Karpuschewski, B., Fast sensor systems for the diagnosis of grinding wheel and workpiece, Proceedings of 5th International Grinding Conference, 1993, MR93-369.
6. Webster, J.A., Dong, W.P. and Lindsay, R., Raw acoustic emission signal analysis of grinding process, Annals of the CIRP, Vol. 45/1/1996, 335-340.
7. Brinksmeier, E., and Werner, F., Monitoring of grinding wheel wear, Annals of the CIRP, Vol. 41/1/1992, 373-376.
8. Oliveira, J.F.G. and Dornfeld, D.A., Dimensional Characterization of Grinding Wheel Surface through Acoustic Emission, Annals of the CIRP, Vol. 43/1/1994, 291-294.

Development of Vibration Cutting using OD-Blade

Ken-ichi ISHIKAWA*, Hitoshi SUWABE* and Michio UNEDA**

*Kanazawa Institute of Technology **Graduate Student

7-1 Ohgigaoka Nonoichi-machi, Ishikawa 921 Japan

1. Introduction

This study aims to develop a new cutting method using OD-Blade¹⁾ with high efficiency, high quality and a long tool life. In this method is applied a low frequency vibration on the workpiece side during the cutting. As shown Figure 1, the vibrational direction is parallel (X-axis)²⁾ and perpendicular (Z-axis) individually to the cutting feed (X-axis) direction. Figure 1 shows a model of the vibration cutting using OD-Blade to Z-axis vibration. In this paper, authors carried out a series of experiments utilizing the effect of vibration. From the results of investigation with regard to the behavior of grinding force, grinding ratio, waviness and chipping for cutting accuracy, this report made the vibration cutting characteristics using OD-Blade clear.

2. Experimental apparatus and method

Figure 2 shows the outline of experimental

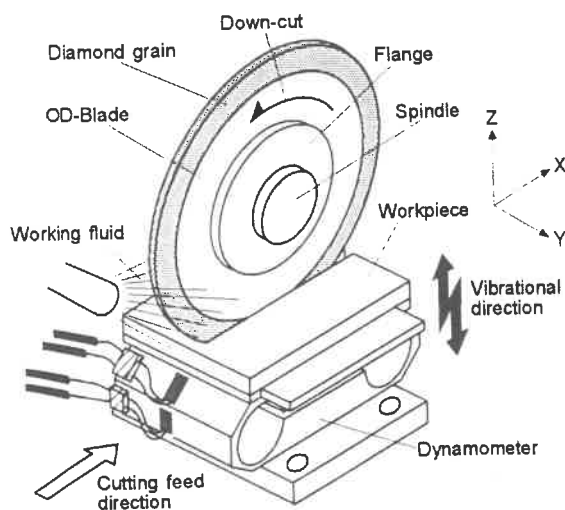


Figure 1 Model of vibration cutting

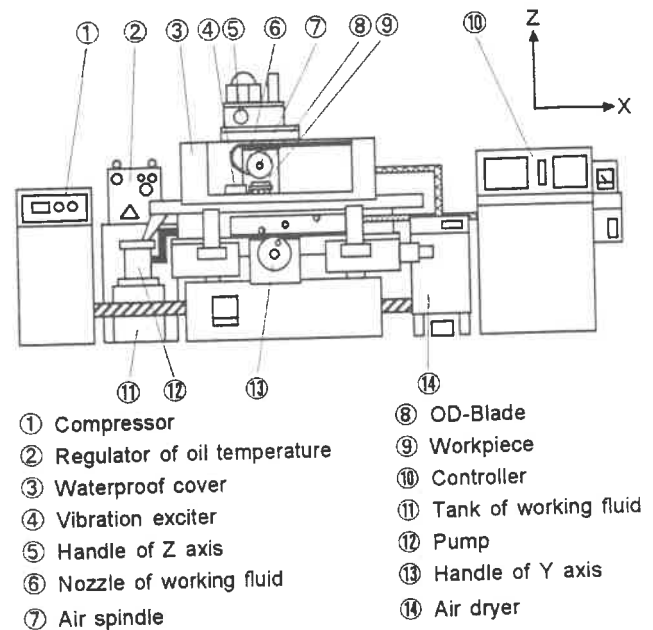


Figure 2 Outline of experimental apparatus

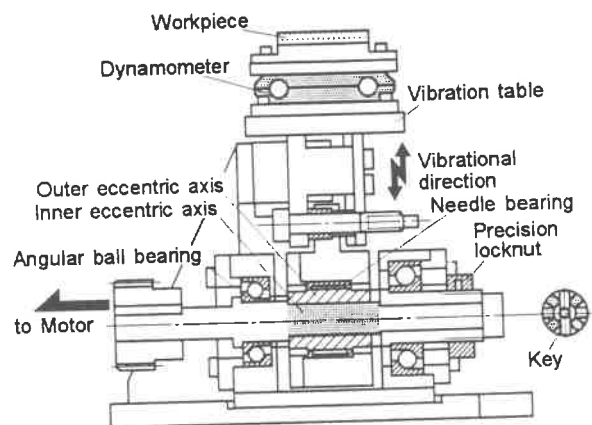


Figure 3 Outline of vibration exciter

apparatus. This apparatus is a slicer using air static pressure bearing in the spindle. The cutting characteristics is evaluated by means of measuring of the grinding force using the octagonal ring dynamometer (natural frequency:325Hz). In the case of the vibration cutting, the vibration exciter for the workpiece as

Table 1 Experimental conditions

OD-Blade	Size	$\phi 125 \times 1\text{mm}$
	Type	Metal-bond, SD#200
Blade speed		1963m/min
Cutting mode		Down-cut
Depth of cut	c	8.0, 11.0mm
Feed speed of workpiece	$\dot{X}$	20 ~ 80mm/min
Frequency	f	0 ~ 20Hz
Amplitude	a	0 ~ 315 μm
Workpiece, Size		Soda glass, $50 \times 60 \times 10\text{mm}$
Grinding fluid type		Water soluble coolant (50:1)
Flow velocity V (Flow rate Q)		223m/min (5328g/min) 358m/min (2203g/min)

shown Figure 3 is set on the working table in the apparatus and is vibrated under the experimental conditions. The double eccentric axes are used for the vibration exciter, and therefore, each amplitude can be operated at 8th stage from 15 ~ 315 μm by means of changing the combination of the inner eccentric axis and the outer eccentric axis using keys at the each axis side face. Also, the applied frequency can be generated from 0 ~ 20Hz by means of adjusting revolution of the speed control motor to the inner eccentric axis through the spur gear. The main experimental conditions are shown in Table 1.

3. Behavior of grinding force

Figure 4 shows the relationship between the applied amplitude and the normal grinding force. The symbols on the axis of ordinates show the conventional cutting and the other's symbols show the vibration cutting. From the experimental results, the normal grinding force in the vibration cutting to each feed speed of the workpiece (20 ~ 80mm/min) shows a decrease compared with one in the conventional cutting. And also, the normal grinding force tends to decrease with increasing the amplitude. It can be explained by a removal model of chip as shown in Figure 5. In the case of the vibration cutting, the processing becomes intermittent cutting owing to cutting and non-cutting phenomena (OD-Blade and the workpiece separate during cutting) in one vibration cycle. In the non-cutting process, inflow of working fluid for the

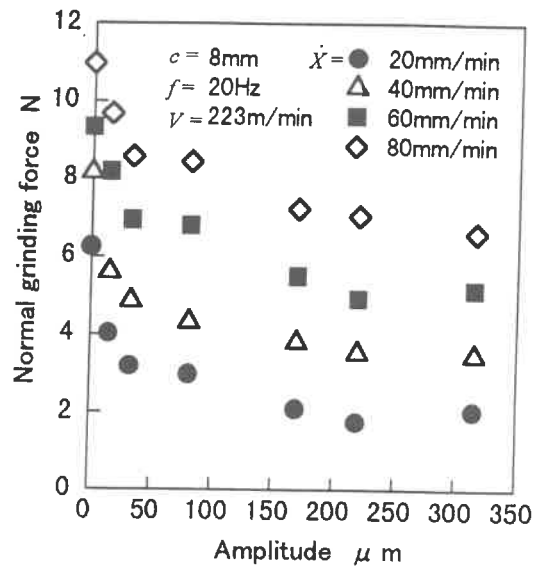


Figure 4 Relationship between amplitude and normal grinding force

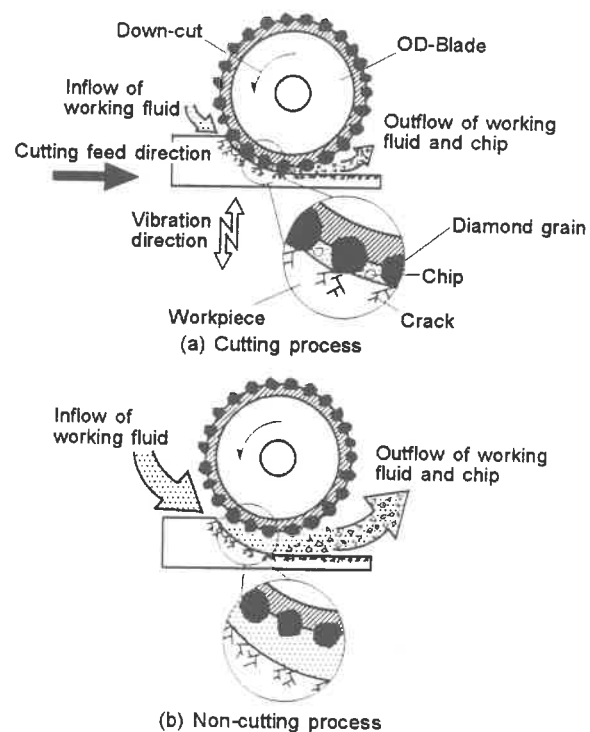


Figure 5 Removal model of chip

processing area is improved. Therefore, the effect of vibration cutting removes chips from the workpiece fully and restrains thermal wear of diamond grains. As a result, the cutting characteristics in the vibration cutting is improved to be compared with one in the

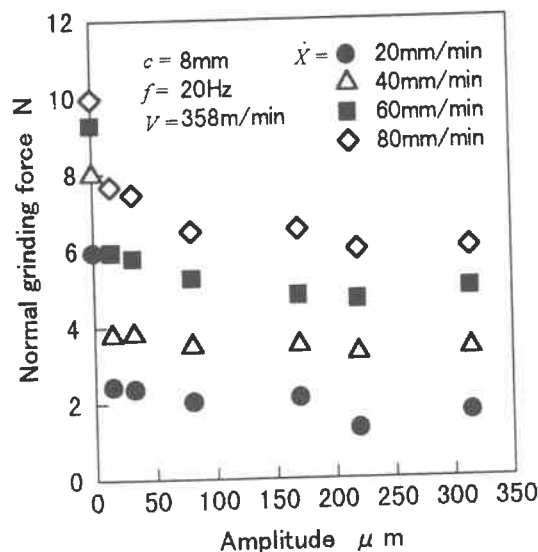


Figure 6 Relationship between amplitude and normal grinding force

conventional cutting, and then the normal grinding force shows a decrease. However, inflow effect of working fluid for the processing area in the case of the lower amplitude is difficult to improve, therefore, a decreasing effect of grinding force is a little. And then, if inflow effect of working fluid can be improved in the case of the lower amplitude, a decreasing effect of grinding force will produce remarkably. Therefore, authors tried an increasing of flow velocity using $\phi 2.8$ pipe as the supplying procedure of working fluid. Figure 6 shows an experimental result in this case. From this result, in the case of the lower amplitude, a decreasing effect of grinding force becomes remarkably in comparison with Figure 4. As a result, it is considered that the cutting characteristics are influenced owing to the supplied method of working fluid in the cutting method using OD-Blade.

4. Dressing effect by vibration

Figure 7 shows the behavior of grinding force in the continuous vibration cutting by means of OD-Blade after non-vibration cutting. From experimental results, in order that a grain

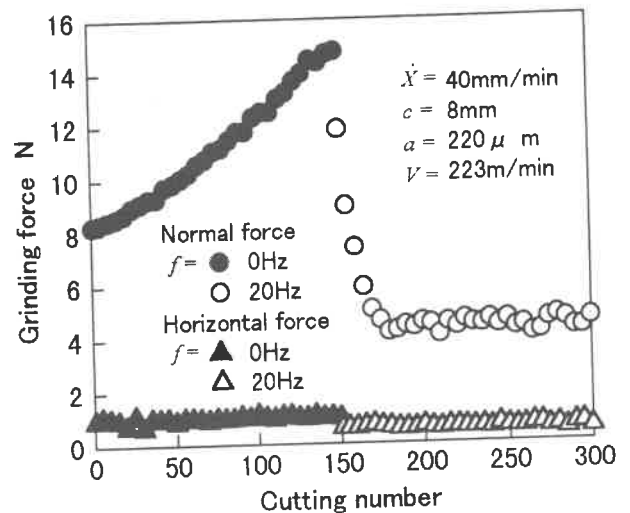


Figure 7 Dressing effect by vibration

is making progress with the attritious wear, loading of chip and breaking down gradually in the non-vibration cutting, the grinding force increases according to the cutting number. However, the grinding force decreases rapidly with vibration cutting within 30 times after the non-vibration cutting and maintains approximately constant value, even if the cutting is successive. Figure 8 shows the observed

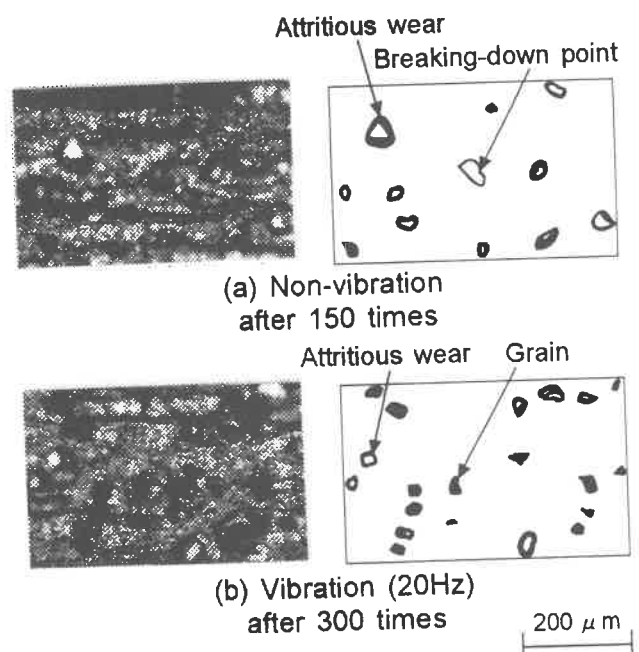


Figure 8 Observation of blade surface by optical microscope

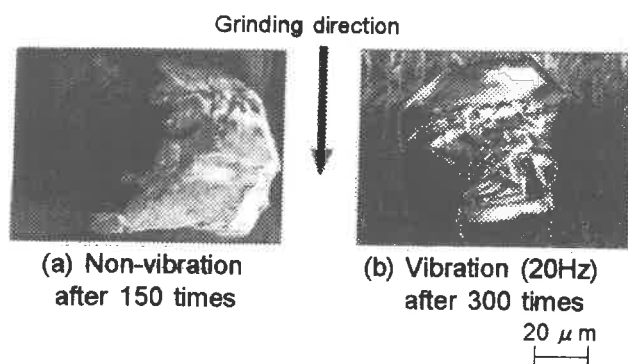


Figure 9 Observation of grain by scanning electron microscope (SEM)

result of OD-Blade surface by means of the optical microscope and Figure 9 shows a grain observed by means of the scanning electron microscope (SEM). In each photograph, (a) shows the non-vibration cutting (after 150 times) and (b) shows the vibration cutting (after 300 times). From these photographs, in the case of the non-vibration cutting, much attritious wear and many breaking down points of grains are observed, and a grain edge becomes roundish owing to attritious wear. However, in the case of the vibration cutting, attritious wear and breaking down points are a little, and much minute cracks have been observed. Therefore, the vibration cutting has the dressing effect of grains and recovers the cutting characteristics even if deteriorated OD-Blade is used.

5. Wear characteristics of OD-Blade

Figure 10 shows the relationship between the applied frequency and wear volume of OD-Blade (Diameter direction, Thickness direction) after 150 times of repetitive cutting. From an experimental result, all wear volume under each condition maintains constant value in the diameter direction. However, wear volume in the vibration cutting decreases as compared with the non-vibration cutting in the thickness direction. In the case of the non-vibration cutting, in general, the elastic deformation of OD-Blade occurs during the cutting process. It is caused

because of irregular volume of grains and shape of OD-Blade. As a result, the side surface of OD-Blade has been brought about attritious wear because it is pushed against the workpiece during the cutting process due to restoring force of OD-Blade. On the other hand, the vibration cutting brings about the separate and contact phenomena between OD-Blade and the workpiece in one vibration cycle. The elastic

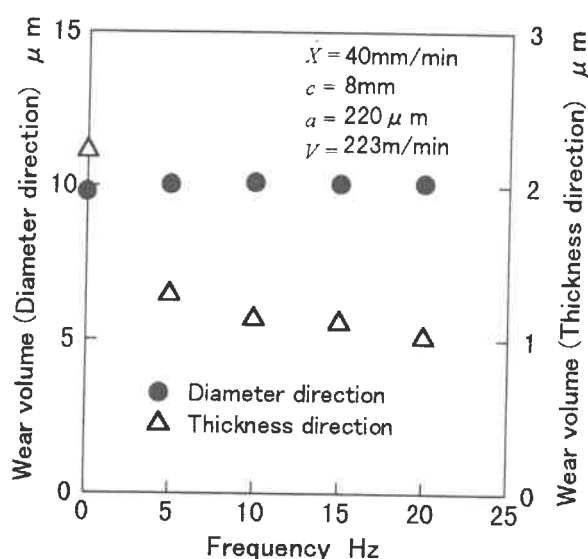


Figure 10 Relationship between frequency and wear volume of OD-Blade

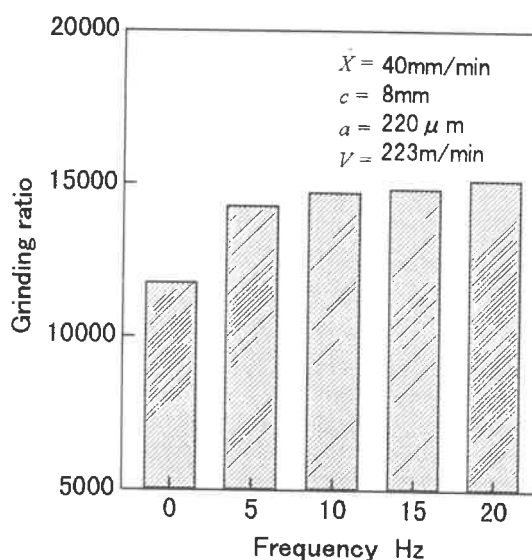


Figure 11 Relationship between frequency and grinding ratio

deformation of OD-Blade restores to the neutral position when the OD-Blade separates from the workpiece due to the elasticity of OD-Blade. Accordingly, the elastic deformation of OD-Blade is controlled in the vibration cutting and wear volume of the side surface of OD-Blade decreases in comparison with the non-vibration cutting. Also, Figure 11 shows the relationship between the applied frequency and the grinding ratio calculated from these results. As shown in the graph, it is clear that the grinding ratio in the vibration cutting increases in comparison with the non-vibration cutting, and the grinding ratio increases with increase of the applied frequency.

6. Cutting accuracy

Figure 12 shows the relationship between the applied frequency and waviness, and Figure 13 shows observed results of chipping formed in the cutting edge. The value C_p shows a quantity of chipping, and C_p is defined as shown in Figure 14. From the experimental results, it is clear that waviness in the vibration cutting is improved as compared with the non-vibration cutting. And also, chipping in the vibration cutting shows smaller than one in the non-vibration

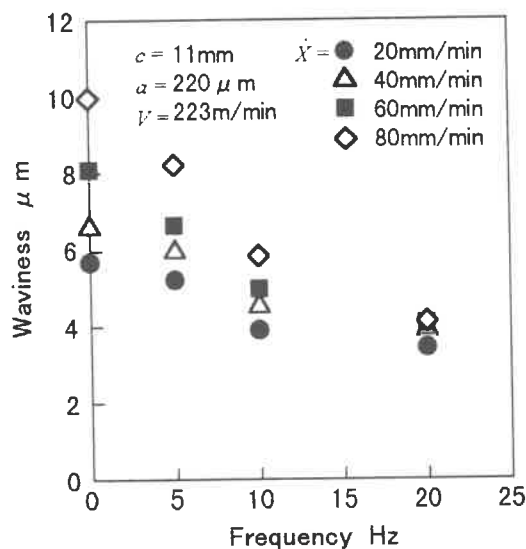
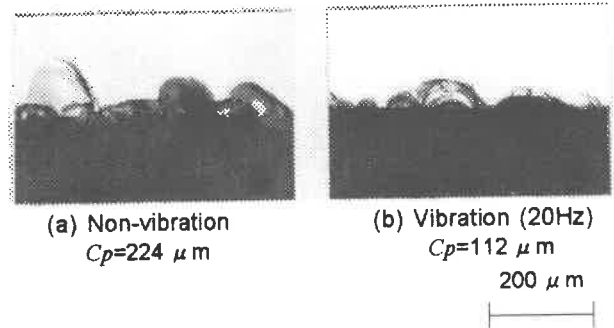
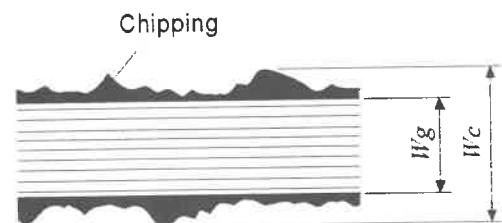


Figure 12 Relationship between frequency and waviness



$\dot{X}=40\text{mm/min}$, $c=8\text{mm}$, $V=223\text{m/min}$
 $f=20\text{Hz}$, $\alpha=33\text{ }\mu\text{m}$

Figure 13 Observation of chipping



W_g : Width of cutting groove

W_c : Width of chipping

C_p : Chipping quantity determinial

$$C_p = \frac{W_c - W_g}{2}$$

Figure 14 Definition of quantitative chipping

cutting. These remarkable results are brought about by means of self corrective effect of OD-Blade during the separate process in one vibration cycle. As a result, in the case of the vibration cutting, the load operated at the cutting edge decreases and chipping in the vibration cutting shows smaller than one in the non-vibration cutting. Also, this self corrective effect of OD-Blade increases according to increasing the applied frequency.

7. Conclusions

Through the experimental results obtained by means of the new cutting method applying low frequency vibration on the workpiece side when the vibrational direction is perpendicular (Z-axis) to the cutting feed (X-axis) direction, the

followings are clear.

- (1)The grinding force in the vibration cutting decreases about 30 ~ 50% in comparison with the non-vibration cutting.
- (2)The vibration cutting brings about a self dressing effect, and the grinding force in the vibration cutting maintains approximately constant value, even if the cutting process is done in succession.
- (3)The grinding ratio in the vibration cutting increases as compared with the non-vibration cutting, and the tool life is improved.
- (4)Waviness in the vibration cutting decreases in comparison with the non-vibration cutting, and chipping in the vibration cutting is smaller than one in the non-vibration cutting.

References

- 1)Ken-ichi ISHIKAWA : High Efficiency and Accuracy Slicing for Hard and Brittle Materials, IPC, (1995).
- 2)Ken-ishi ISHIKAWA, Hitoshi SUWABE, Kazuya MASUDA and Michio UNEDA : A Study on Vibration Cutting Characteristics Using OD-Blade, Journal of the Japan Society for Precision Engineering, (1996), 62, 3, pp.438 ~ 442.

PROSPECTIVE OF FRACTAL APPLICATION IN GRINDING

X. M. Liu, G. Q. Cai, H. W. Zheng,

Mechanical Engineering Institute, Northeastern University P. R. China

D. L. Ma

Liaoning Industrial Art School P. R. China

Q. Y. Wen

Shenyang Donghang Mechanical & Electrical Technology Institute P. R. China

ABSTRACT

In this paper the concept of fractal which is a new mathematics branch developed in recent year is briefly introduced. Then, the dimension definitions and calculation methods are stated. At last, the prospective about fractal application in grinding is given.

KEY WORDS: Fractal Application Grinding

1. INTRODUCTION

Many phenomena in nature such as shape of cloud, rugged mountain and curve of coastline, etc., often show an irregular and altered state which can not be well described in Euclidean Space.

Mathematicians have developed geometrical concepts that transcend traditional geometrical concepts, but unfortunately, the concepts have failed in the past to gain acceptance in the natural sciences because of the rather abstract and "pedantic" presentation. Traditionally, the Euclidean lines, circles, spheres and tetrahedral have served as basis of the intuitive understanding of nature. Such shape-primitive representation function well in man-made carpentered environments. The problem, then, is how shall we describe the shape of a crumpled newspaper? a clump of leaves? a jagged mountain? When we attempt to describe the crumpled surface typical of natural objects, however, the result is usually complex. Such awkwardness makes these shape-primitive representations difficult to envisage as the basis for human performance level capabilities.

During the last 30 years, B. B. Mandelbrot has developed and popularized a relatively novel class of mathematical functions known as fractal. It has been proved that many phenomena and processes in nature have fractal structure.

The defining characteristic of a fractal is that it has a fractional dimension from which we get the word "fractal". Technically, a fractal is defined as a set for which the Hausdorff-Besicovich dimension is strictly larger than the Topological dimension. The point is that the definition although correct and precise, is too restrictive. Mandelbrot (1982) has retracted this tentative definition and proposes instead the following: A fractal is a shape made of parts similar to the whole in some way. A neat and complete characterization of fractal is still lacking (Mandelbrot, 1987).

The triadic Koch curve is an example which is used to illustrate that a curve may have a fractal dimension $D > 1$ shown in Fig. 1.

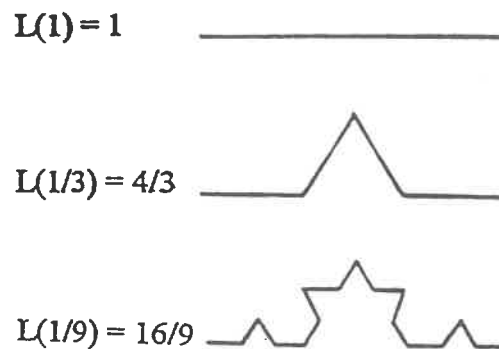


Fig. 1 Generation of Koch Curve

The picture illustrates an iterative or recursive procedure for constructing a fractal curve. A simple line segment is divided into thirds and the middle segment is replaced by two equal segments forming part of an equilateral triangle. At the next stage in the construction each of these 4 segments is replaced by 4 new segments with length $1/3$ of their parent according to the original pattern. This procedure repeated over and over, will yield the beautiful Von Koch curve.

Shapes in world can be divided into two kinds:

- (a) ones which have characteristic length and are smooth and differential;
- (b) the others which have no characteristic length and are not smooth and differential.

Fractal geometry is mainly to deal with the later ones which include self-similar curves, positions and shapes, self-reverse irregular shapes, self-squaring transform, self-affinity transform and set. The essential characteristic of them is the scaling invariance. At present, the researches are mainly ranged in things which have self-similarity which can be popularly interpreted as the characteristic structure reappears at different scales, i.e., multiplying any parts is statistically same as whole.

Fractal dimension is a kind of measurement. The larger the dimension is, the more complex the structure is, e.g., an object with dimension of $0 < D < 1$ describing an object with 0 length and ∞ points (Cantor Set), $1 < D < 2$ describing an object with 0 area and ∞ length (Serpinsky Gasket), $2 < D < 3$ describing an object with 0 volume and ∞ area (Serpinsky Sponge).

2. DIMENSION DEFINITION

At present there have not been a strict and overall definition for fractal dimension. The fractal concept is still in developing. Different definition determines different calculation method. Here the three commonly used fractal dimension are briefly introduced as below.

1) Box Dimension

How do we measure the "size" of a set S of points in space? A simple way to measure the length of curve, the area of surface or the volume of an object is to divide space into small cubes of side δ as an example illustrated in Fig. 2.



Fig. 2 Measuring the Size of a Curve

Here we may use small spheres of diameter δ instead. If we center a small sphere on a point in the set then all the points that are at a distance $r < 1/2 \delta$ from the point at the center are covered by the sphere. By counting the number of spheres needed to cover the set of points we obtain a measure of the size of the set. So far, in order to give a measure of the size of a set of points in space we have taken a test function $h(\delta) = \gamma(d)\delta^d$ - a line, square, disk, sphere or cube - and have covered the set to form the measure $M_d = \sum h(\delta)$. Here $\gamma(d)$ is geometrical factor. In general we find that as $\delta \rightarrow 0$, the measure M_d is either zero or infinite depending on the choice of d - the dimension of the measure. Then we call M_d the d - measure of the set. If $N(\delta)$ is the numbers of the small spheres with diameter δ , then the Box Dimension D_0 can be defined as:

$$D_0 = - \lim_{\delta \rightarrow 0} \ln N(\delta) / \ln \delta \quad (1)$$

Let us take coastline dimension as an example. On a map of coast, we could walk a divider with η an opening corresponding to δ km along the coastline on the map and count the numbers of steps $N(\delta)$ needed to move from one end of the map to the other. At beginning, we would choose such a large opening of the diameter that we would not have to bother about even the deepest fjord and estimate the length to be $L(\delta) = N(\delta)\delta$. Then, we would use a somewhat smaller opening δ and try again. This time the large fjords would contribute to the measured length. Every time we increased the resolution we find an increase in the measured length of the coastline. In fact, the measured length is nicely approximated by the formula $L(\delta) = a\delta^{1-D}$. Here, the a is constant and D is fractal dimension. So the coastline is fractal with a fractal dimension D . We may determine the fractal dimension of the coastline by finding the slope of $\ln N(\delta)$, plotted as function of $\ln \delta$. The slope of the line in $\ln$ - $\ln$ plot is $1-D$. The dimension determined from equation (1) by counting the number of spheres needed to cover the set as a function of the sphere size is now called the Box Dimension.

2) Information Dimension

In calculation of Box Dimension we count only the number of δ spheres and do not consider the cover points of each sphere. If it is considered the points density of δ spheres is different. Let P_i be the point density, then the $P_i \ln P_i$ shows the entropy of the systems, i.e., the information providing ability of the system. The information dimension D_1 can be defined by the formula:

$$D_1 = - \lim_{\delta \rightarrow 0} \left[\sum_{i=1}^N P_i \ln(1/P_i) \right] / \ln \delta \quad (2)$$

This expression is same as formula (1) when the point density of spheres is even. That is when $P_i = 1/N(\delta)$, the $D_1 = D_0$. As stated above the expression shows the relationship between entropy of system to fractal dimension.

3) Relation Dimension

P. Grassberger and I. Procaccia gave the definition of relation dimension D_2 as below:

$$D_2 = -\lim_{\delta \rightarrow 0} \text{Ln}C(\delta) / \text{Ln}\delta \quad (3)$$

Where the $C(\delta)$ is a series of system results. If the δ is properly selected, there will be a scale range in which there is a relation $C(\delta) = \delta^D$. Here the D is approximate value of D_2 . In calculation of relation dimension the value of δ must be selected in a proper range, because if δ is too big, then $C(\delta) = 1$ and $\text{Ln}C(\delta) = 0$, thus the δ may not represent inner nature of system and if δ is too small then some accidental events may affect the real nature of system.

We briefly introduce the three fractal dimensions above, what is the relationship among D_0 , D_1 , D_2 ? It can be theoretically proved that there are unlimited kinds of dimensions, which can be stated as below:

$$D_q = 1 / (q-1) \lim_{\delta \rightarrow 0} (\text{Ln} \sum_i P_i^q) / \text{Ln}\delta \quad (q = 0, 1, \dots, n) \quad (4)$$

In this formula when $q = 0, 1$, and 2 the D_q will be D_0 , D_1 and D_2 . They obey the rule $D_q > D_p$ (when $p > q$), i.e., $D_T \leq D_q \leq \dots \leq D_2 \leq D_1 \leq D_0$ as shown in Fig.3.

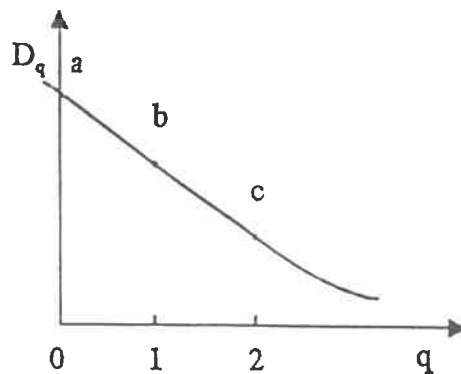


Fig.3 Relationship Curve between D_q and q

3. APPLICATION PROSPECTIVE IN GRINDING

Although the fractal geometry has been applied in many engineering, the authors until now have not found any application in grinding field. From the view of research object of fractal and characteristics of grinding, it is possible to find some applications of fractal in the field, which can be explained as below:

(a) Fractal is effective in dealing with objects which have irregular or scattered characteristics;

(b) Grinding is one of the process which have strong irregular and scattered characteristics.

First, considering from the grinding tools the grains which form surface of wheel, belt, etc., are irregular particles and so the surface of these tools are also rugged. The surface geometry of wheels, belts, and so on, which has great influence on grinding parameters has been an important research aspect in grinding for a long time. The researches in past have been strongly based on Plato Shape, this simplifying leads to inexact description and losing many important information about the real nature. While the fractal geometry may be effective to deal with the problem for its characteristics.

Second, from the aspect of machining the process of tool wear and surface generation of workpiece has strong scattered characteristics especially in machining fragile materials such as ceramics, stone, and so on. The whole process consists of many small and random events. The irregular surface is affected by load and inner micro structure of materials. In this situation, micro fracture play an important role. There may be a number of cases in which microscopic fracture mechanism can create fractal characteristic in a certain range of scales.

From the view of mathematics the irregularity belong to nature of geometry and the fractal geometry may be found some usages in dealing with those such as the relationship between dimension and surface roughness of workpiece, etc.

Moreover, the fractal geometry can be used in computer technology to simulate and study irregular shape and roughness of workpiece, wear process, etc.

At present there have been some applications in rock crushing researches. The results are successful. Mandelbrot pointed that the rock surface and distribution of crushed pieces have unchanged scale or are fractal. Turcotto [3] proved that the crushed pieces of rock caused by explosion and impaction are fractal in structure, e.g., the particle distribution of rock crushed by lead hammer has fractal dimension of $D = 1.44$ [4]. The distribution of celestial body has also been proved to be fractal in structure.

In short, fractal is one kind of expression of nature. In order to know nature with fractal it is necessary to set up models and give analysis method. So that the essential relation between surface and true nature can be revealed. Because the fractal is a new branch of mathematics, it need to be more developed both from itself and application. Fractal is still at an early stage of development. But it is sure that with the development and completion of fractal geometry it will become one of powerful means to quantitatively know the objective world.

REFERENCE:

- [1] B. B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, 1983
- [2] B. B. Mandelbrot, Stochastic Model for the Earth's Relief, the Shape and Fractal Dimension of the Coastline, *Proc. Natl. Acad. Sci USA*, 72, 3825-3828, 1975
- [3] T. L. Turcotto, A Fractal Model for Crustal Deformation, *Tectonophys*, 132, 261-269, 1987
- [4] M. A. Lange et al, Impact Gratering and Spall Failure of Gabbro, *Icarus*, 58, 383-395, 1984
- [5] L. Z. Fong, Fractal of Large Scale Structure in the Universe, *Modern Physics Letter A*, 1, 11, 601-605, 1986

Grinding Characteristics of Plasma-Sprayed Ceramics

K. Narayanasamy

Associate Professor

Dept. of Mechanical Engineering

Indian Institute of Technology

Madras - 600 036, India.

S. Gowri

Senior Lecturer

Dept. of Mechanical Engineering

Karnataka Regional Engineering College

Srinivasnagar - 574 157, India.

Rapid developments in ceramic materials and their processing techniques have widened the scope of applications either as monolithic or sprayed coatings. In any case, in order to meet the desired functions, one has to maintain the requisite surface finish, bearing area fraction and dimensional tolerance. These can be achieved by precision machining of ceramics by way of grinding, lapping and superfinishing. Since ceramics have low fracture/rupture strength, it is necessary to take utmost care during precision machining of ceramics, in order to avoid the occurrence of surface cracks and other associated defects. This paper presents an account on the grinding of Al_2O_3 and partially stabilized zirconia (PSZ) coatings.

Experimental:

Plasma sprayed Al_2O_3 and PSZ coatings were surface ground using diamond and CBN. Grinding performance was evaluated in terms of grinding force and surface finish. Grinding force was measured using piezo electric type Kistler dynamometer. In order to assess the nature of grinding, Acoustic emission signals from the work piece were sensed using AET 5500 Acoustic emission recording system.

Parametric Influence on Cutting force

During grinding of plasma sprayed ceramics, the grinding force components were measured using piezoelectric dynamometer. Typical variations of normal force of grinding for Al_2O_3 and PSZ ceramics are shown in Fig.1. The grinding force is influenced by the cutting conditions, work material and the type of abrasives being used for grinding. It is seen that as the grinding speed increases, there is a reduction in grinding force in the case of diamond grinding, while raising trend can be seen in the case of grinding with CBN. Ploughing of diamond on the work surface with lower grinding velocity results in higher order F_N , while grinding performance at high speeds results in reduced order of grinding force. CBN grinds the ceramics with relatively higher order force. This can be attributed to the occurrence of grain-flattening. Both Al_2O_3 and PSZ exhibited certain phase transformation during grinding as illustrated in Table.1.

Table 1 Influence of grinding velocity on phase content of PSZ / Al_2O_3

Velocity m/s	PSZ m/t ratio Diamond	PSZ m/t ratio CBN	Al_2O_3 I _R ratio Diamond	Al_2O_3 I _R ratio CBN
5.7	0.153	0.150	0.202	0.200
11.5	0.166	0.159	0.164	0.153
15.7	0.194	0.159	0.140	0.142
31.5	0.194	0.162	0.138	0.141

Transformation of alpha to gamma Al_2O_3 during grinding reduces the hardness and consequently a reduction in the grinding force F_N . However, in the case of PSZ the $t \rightarrow m$ transformation increases the toughness and hence results in higher order F_N . During grinding, PSZ ceramic experienced mechanical stress and rise in temperature. Depending up on the temperature or stress magnitude, it was possible that PSZ surface could undergo stress-induced $t \rightarrow m$ transformation or with higher order temperature, it may experience $t \rightarrow m \rightarrow t$ transformations [3]. At low speed range, due to mechanical plowing, the dominant mechanical stress might be resulting in $t \rightarrow m$ transformation promoting transformation toughening of surface and hence the normal force increases.

Process status identification through AE parameters

Ceramics are brittle materials, known for the presence of isolated strength pockets over their matrix; hence during grinding of ceramics, depending upon the intensity of applied grinding stress and strength of the materials in the grinding zone, material removal would be associated with either deformation induced shear or clear rupture of the material [1]. When an abrasive enters the ceramic surface the material would be deformed resulting in flow of material around the intending abrasive. During indentation, the ceramic experiences median and lateral cracking in the material over the vicinity of the tip of the indenter as illustrated in Fig.2. In order to eliminate cracking, one has to limit the force of indentation. The permissible force of indentation without occurrence of cracking can be stated as [2]:

$$p = \frac{(K_{ic})^4 \cdot A (E/H)}{H^3}$$

K_{ic} = fracture toughness
 H = hardness

Since the fracture toughness K_{ic} of ceramics is of lower magnitude, one has to control/limit the depth of grinding for defect free grinding. Limiting the depth of grinding produces ductile grinding of ceramics associated with deformation and shearing of ceramics and good surface texture.

During grinding of ceramics, the material can be removed by plastic deformation and shearing or instantaneous rupturing depending on the applied stress. The nature of material removal can be studied by sensing the AE signal emanating from the ceramics. AE signal parameters such as Event Duration (ED) and Ringdown Counts (RDC) were sensed during this study. Typical observations are presented in Fig.3. The steady reduction in the grinding force in the case of diamond grinding of Al_2O_3 has resulted in deformation and shearing of surface material as shown by the raising trend in ED and RDC. The relatively lower magnitude of ED and RDC in the case of CBN grinding is due to higher order grinding force F_N inducing rupturing of surface material. The near ductile regime of grinding due to lower order force exhibited by diamond grinding of ceramics results in production of better surface texture of ground ceramics.

Observation on Surface Finish

Typical observations on surface finish are presented in Fig.4. It is seen that with the grinding speed, there is an improvement in surface finish for both Al_2O_3 and PSZ ceramics. Referring to the figure, It can be seen that marginally rougher texture was observed in the case of CBN grinding. Apart from surface finish, the bearing area characteristic of the ground surfaces were also recorded as presented in Fig.5. Referring to the figure, there is a distinct difference in the bearing area characteristics of ceramics in the as-sprayed condition, especially the saturation tendency in the range of 80-100% t_p in the case of PSZ can be attributed to the occurrence of $t \rightarrow m$ transformation of the material on striking the surface

associated with grain-flattening. During grinding, Al_2O_3 exhibits faster bearing area improvements, which can be attributed to the relatively higher hardness. Among the abrasives, CBN grinding results in a texture, exhibiting rapid bearing area development. The deformation of asperities over the ceramic surface due to grain-flattening of CBN abrasive may be the possible cause.

Conclusion

1. While grinding of ceramic coatings, grinding velocity, range of 11.5 to 16 m/s, and depth of grinding of 30 μm were evaluated to be critical.
2. Diamond wheel grinding, compared to that of CBN, gave better surface finish.
3. In case of Al_2O_3 , the surface texture with faster bearing area improvement was observed.
4. PSZ ground by diamond showed higher $t \rightarrow m$ phase transformation while it was limited in case of CBN grinding because of the relatively "cool" grinding of CBN.
5. AE parameters such as event duration and ring down count indicated good grinding associated with rupture of surface material when ground at 10-15 m/s speed; also indicated the occurrence of abrasive-dullness and consequent changes in the nature of grinding from perfect rupturing to one associated with deformation and shearing (ductile grinding).

Key words : plasma spraying, ceramics, grinding process, phase transformation, acoustic emission, surface finish.

References

1. Namba, Y., Yamada, Y. Tsuboi, A., Unno, K., and Nakao, H., "Structure of Mn-Zn ferrite single crystals ground by an Ultraprecision surface grinding with various diamond wheels", *Annals of CIRP*, 1992, Vol.41, No.1, pp.347-351.
2. Krishnamurthy, R., Arunachalam, L.M., and Gokularathinam, C.V., "Grinding of transformation-toughened Y-TZP ceramics", *Annals of CIRP*, 1991, Vol.40, No.1, pp.331-337.
3. Miller, R.A., Smilake, J.L., and Garlick, R.G., "Science and Technology of ZrO_2 -Advances in ceramics", *Journal of American Ceramic Society*, 1981, pp.241-253.

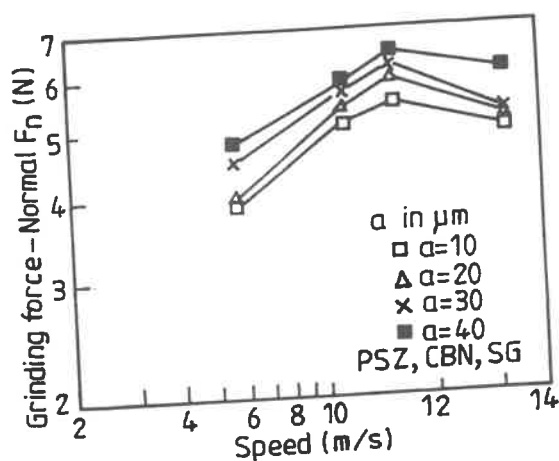
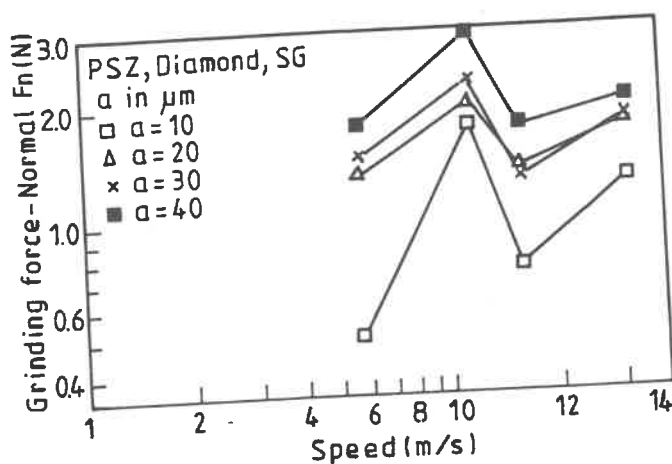
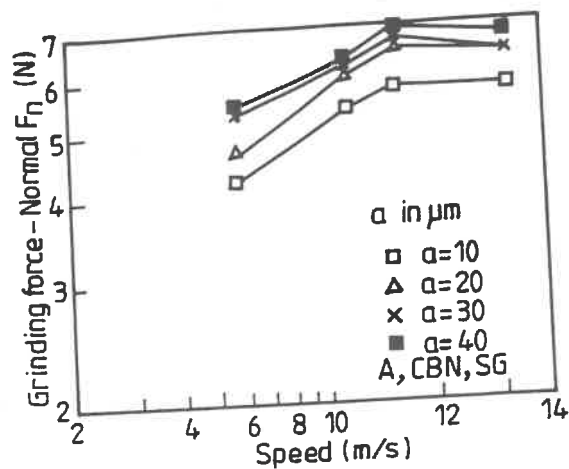
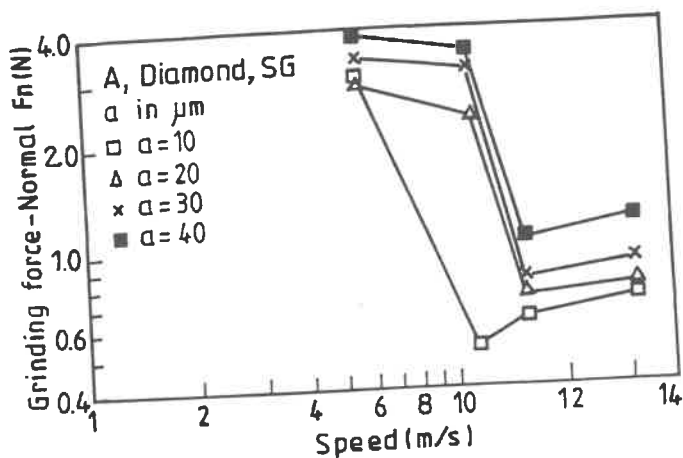


Fig.1

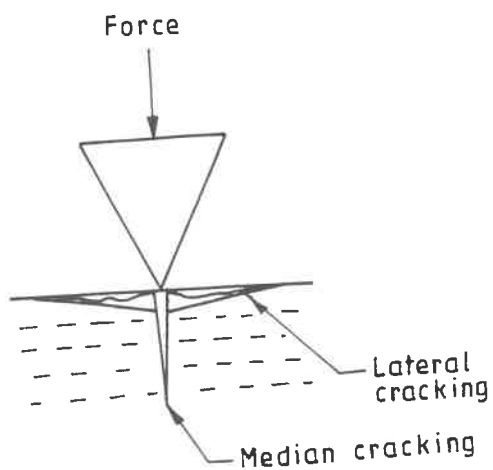


Fig.2

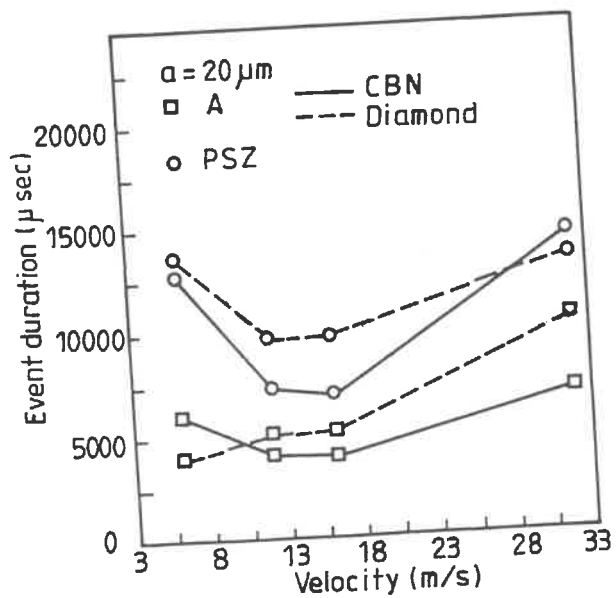


Fig. 3

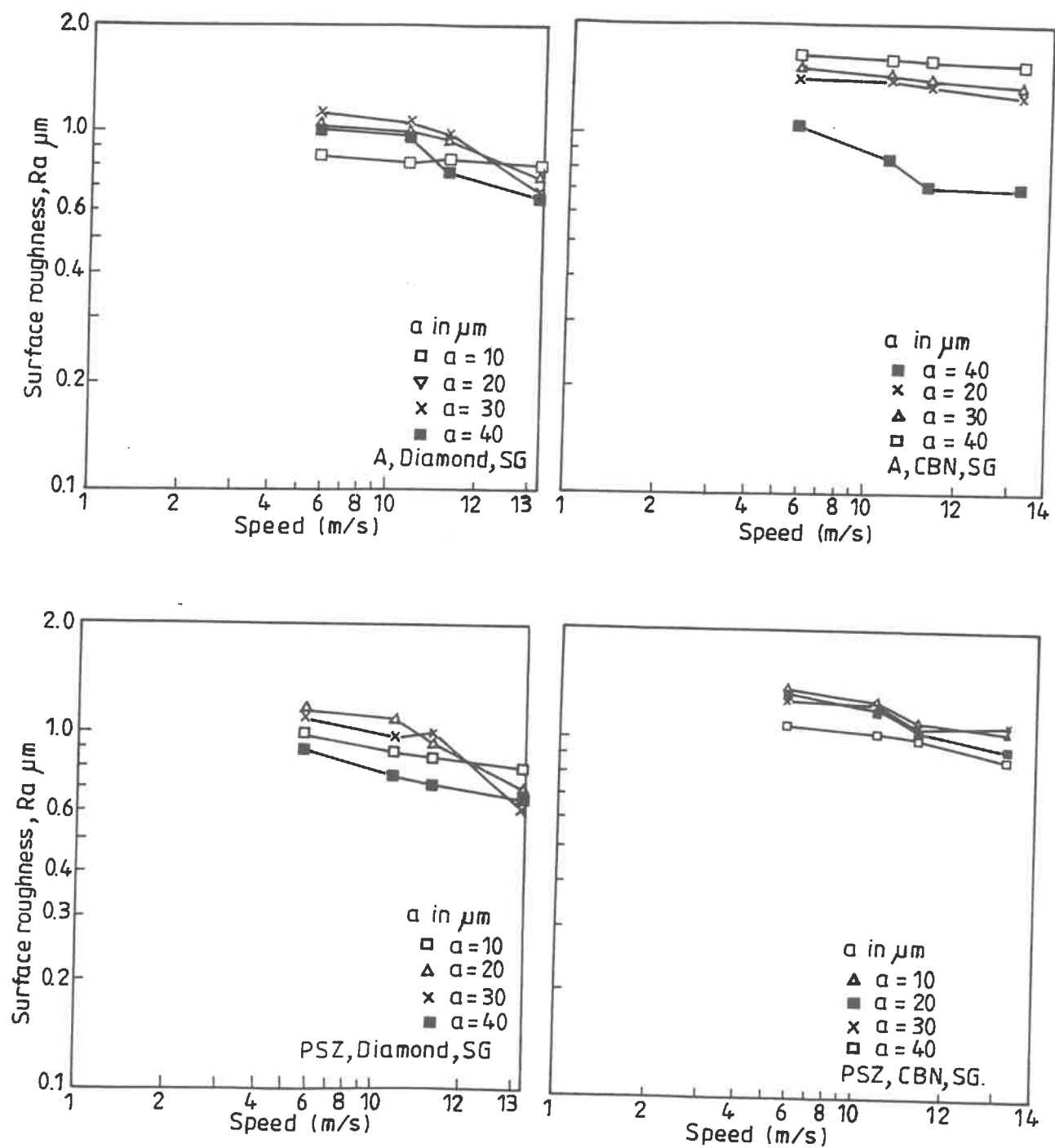


Fig. 4

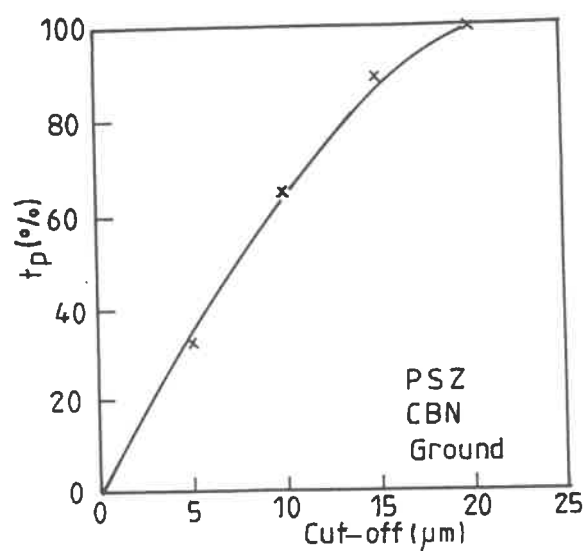
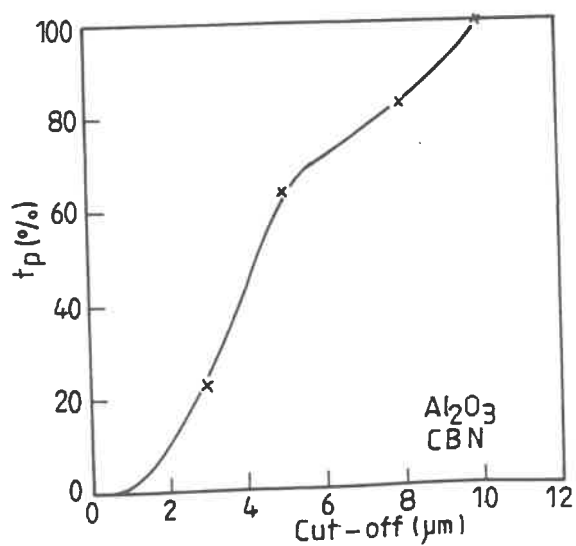
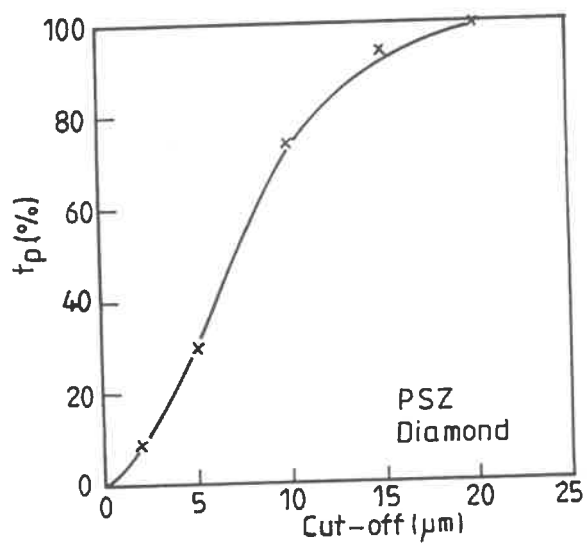
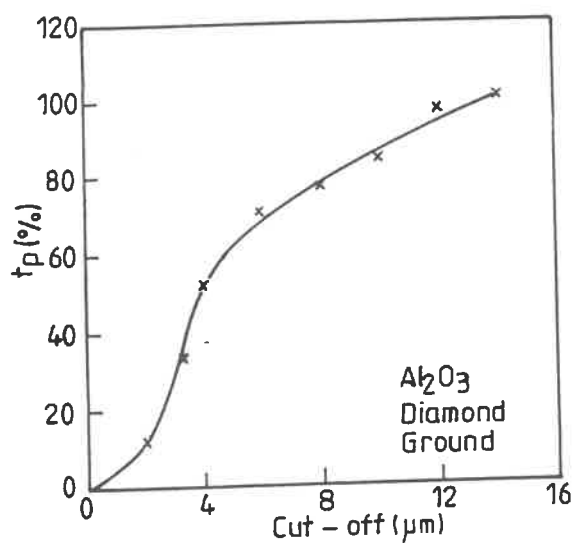
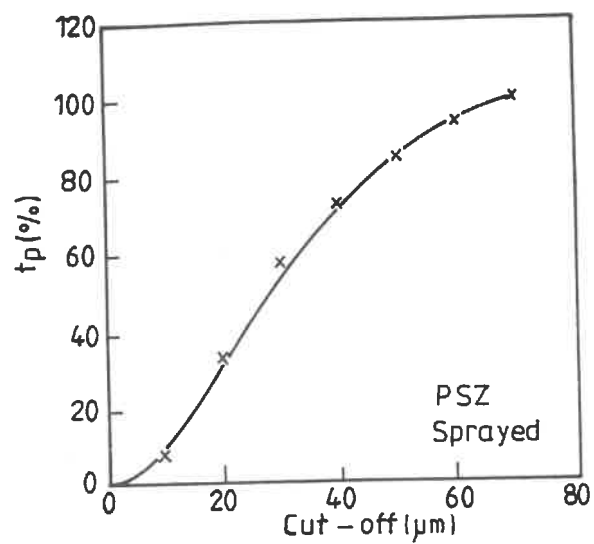
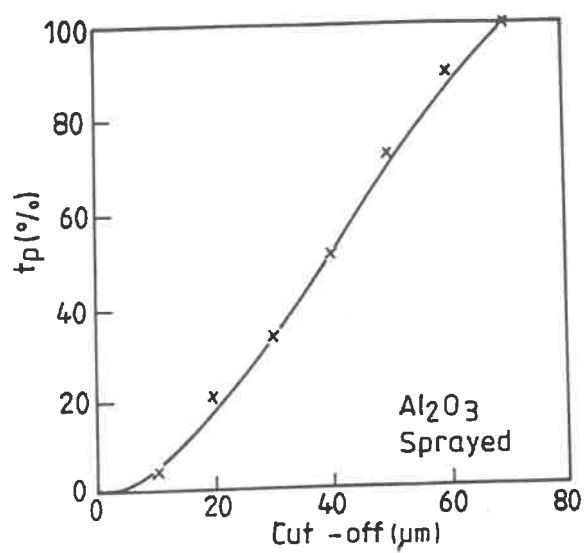


Fig. 5

THE EFFECT OF WORKPIECE MODULATION ON GRINDING FORCE

Weimin Qu, Michele H. Miller
Department of ME-EM
Michigan Technological University
Houghton, MI

1 Introduction

High bandwidth actuators have made possible or improved performance in many precision engineering applications, such as high resolution microscopes and fine position control of structures. Another potential area of application is improvement of the machining process. Ultrasonic ($>20,000$ Hz) vibrations have been used to inhibit chatter in metal cutting [1] in addition to reducing fracture in the single point cutting [2] and grinding [3] of brittle materials. Surface roughness has been improved by modulating the workpiece transverse to the cutting direction at 1000 Hz [4]. The reasons for improvement are not well understood in all cases but are postulated to involve geometry effects, friction effects, and material behavior effects. Simulation has shown that geometric effects can adequately explain the improvement in surface roughness [4] but not the reduction of fracture [5] or inhibition of chatter. The present paper focuses on the effect of high frequency modulation on grinding force.

2 Theoretical Concept

One of the effects of high frequency relative motion between two materials is to reduce the effective friction force [6,7]. In machining, frictional forces arise due to relative motion between the chip and tool rake face and between the machined surface and tool flank face. A reduction in these frictional forces reduces the force components in the cutting and thrust directions. A lower cutting force is beneficial in that it reduces the power required while a lower thrust force reduces the possibility of chatter or fracture.

Moriwaki [2] achieved a seven times increase in critical depth of cut for single point machining of glass; however, the cutting speed was quite low--1.5 m/min. With the much higher cutting speed of grinding and the short time of each abrasive/work interaction, it is uncertain whether high frequency vibration would have any effect on abrasive/work friction and, consequently, performance.

3 Description of Experiments

Experiments were conducted to assess the effect of modulation on grinding forces. Figure 1 shows a schematic illustration of a surface grinder with a modulating workpiece holder attached to its feed table. The workpiece holder is driven in the crossfeed (or X) direction by a magnetostrictive actuator whose nominal maximum vibration frequency is 5000 Hz. A three axis dynamometer provides force measurement.

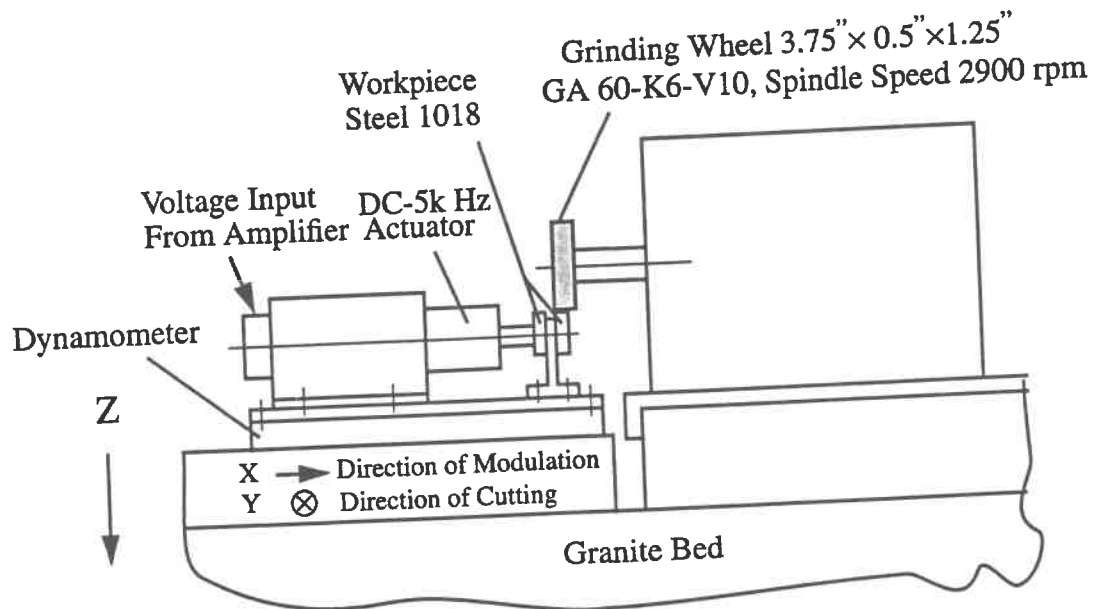


Figure 1: Schematic illustration of surface grinder equipped with actuator and force dynamometer

Cutting tests were performed at four modulation frequencies and with no modulation. The input amplitude of the sinusoidal voltage signal into the actuator amplifier was held constant at 125 mV for all frequencies. The output displacement amplitude was between 4 and 4.5 μm for three of the four frequencies and 9.5 μm for the resonant frequency (4000 Hz). Table 1 summarizes the output displacements. Significant coupling to the Y and Z directions occurs which may impact the force results. Experimental parameters common to all tests are listed in Table 2.

Frequency (Hz)	Input Signal Voltage (mV)	Output Displacement (μm)		
		X	Y	Z
1000	125	4	0	0
2500	125	4.5	0	2.1
3000	125	4.25	0	2.5
4000	125	9.5	2.12	3.6
4600	125	4.25	0	1.4

Table 1: Measurement of workpiece vibration

Figures 2(a) and (b) show examples of force measurements for the thrust direction (Z) without and with 1 kHz modulation. With modulation the thrust force decreases by 4.8% while the force in the cutting direction (not shown) decreases by 6.8%. The plotted forces have been averaged to remove the high frequency components. The raw force signal for the test with modulation included the inertial forces of the actuator system moving on top of the dynamometer. The inertial force occurs at the modulation frequency; its mean value is zero (which can be seen when not grinding). The inertial force is not apparent in Figures 2(a) and (b) due to the averaging.

Spindle Speed	2900 rpm
Cutting Speed	14.5 m/s
Feed Rate	19 mm/s
Depth of Cut	0.025 mm
Workpiece	1018 Low Carbon Steel

Table 2: Operating conditions for grinding experiments

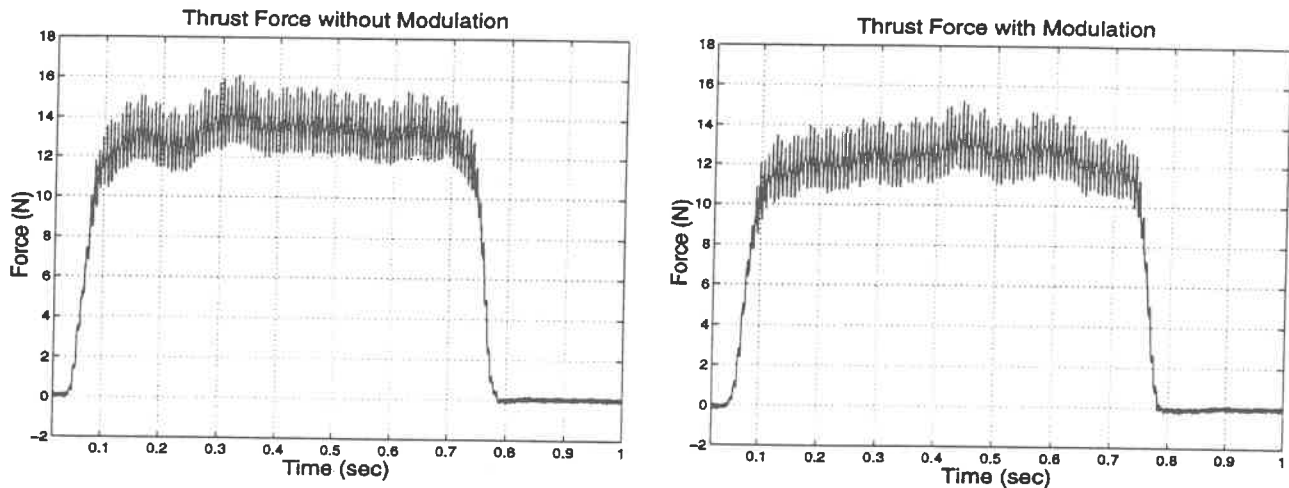


Figure 2: Grinding thrust force signals (a) without modulation and (b) with 1 kHz modulation. Data has been averaged to remove inertial force component.

Tests at higher frequencies exhibit a higher reduction in force as shown in Figures 3(a) and (b). Each data point in these figures represents the average of two tests. The highest reduction in force occurs at the resonant frequency. For this case, however, the vibration amplitude was significantly higher than for the other frequencies. Therefore, in addition to frequency, amplitude may have an important effect on the results and needs to be studied further.

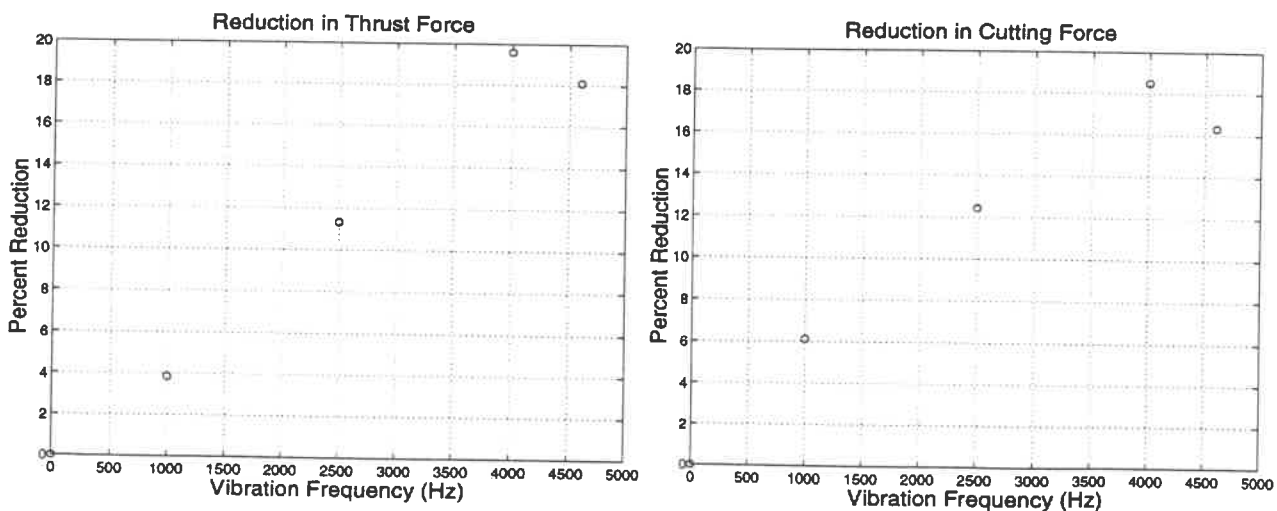


Figure 3: Reduction in (a) thrust and (b) cutting forces as a function of modulation frequency

4 Summary and Conclusions

Experimental studies of the modulation grinding of 1018 low carbon steel have demonstrated the following:

- (1) High frequency modulation transverse to the cutting direction reduces the cutting and thrust forces in surface grinding.
- (2) At the resonant frequency, the effect of modulation on forces is the biggest. The cutting force changes by 16.6% and the thrust force by 18.1%.
- (3) An increase of the amplitude and frequency of modulation is beneficial to the reduction of the cutting and thrust force. Tests at even higher frequencies are required to find optimum values for amplitude and frequency.

Similar tests will be performed on brittle workpiece materials to determine if they also can be ground with less force and if the reduction in force reduces surface fracture.

References

- [1] Shey, J. A., *Tribology in Metalworking: Friction, Lubrication, and Wear*, Amer. Soc. for Metals, Metal Park, OH, 1983.
- [2] Moriwaki, T., Shamoto, E. and Inoue, K., "Ultraprecision Ductile Cutting of Glass by Applying Ultrasonic Vibration," *Annals of the CIRP*, Vol. 41, No. 1, pp. 141-144, 1992.
- [3] Khanna, N., Pei, Z. J. and Ferreira, P. M., "An Experimental Investigation of Rotary Ultrasonic Grinding of Ceramic Disks," *Technical Papers of the North American Manufacturing Research Institution of SME*, pp. 67-72, 1995.
- [4] Wang, Y., Moon, K. S. and Miller, M. H., "Micro-Actuation in Precision Grinding," *Proceedings of the 1996 ASME IMECE*.
- [5] Miller, M. H. and Moon, K. S., "The Effect of Workpiece Modulation on Grinding Kinematics," *Proceedings of the ASPE Spring Topical Meeting on Precision Grinding of Brittle Materials*, pp. 70-75, 1996.
- [6] Mohri, N., and Saito, N., "Some Effects of Ultrasonic Vibration on the Inserting Operation," *Int J Adv Manuf Technol*, Vol 9, pp. 225-230, 1994.
- [7] Moriwaki, T. and Shamoto, E., "Ultrasonic Elliptical Vibration Cutting," *Annals of the CIRP*, Vol. 44, No. 1, pp. 31-34, 1995.

Acknowledgments

The authors wish to acknowledge the financial support provided by NSF Grant No. DMI9521494 and by the SME Education Foundation. The authors also thank the NSF/DARPA MT-AMRI for use of the magnetostrictive actuator.

DEVELOPMENT OF GRINDING MODELS FOR BRITTLE MATERIALS

Keith Sharp, Ronald O. Scattergood
Precision Engineering Center
North Carolina State University
Raleigh, NC

INTRODUCTION

The fracture mechanics used to predict subsurface damage during machining requires knowledge of the depth of cut of the tool. For brittle materials, analysis shows that there will be ductile response for cutting depths below a critical value and brittle response for depths above this value. Finding the depth of cut is relatively straightforward for single-point machining such as diamond turning, but much more difficult for the large number of randomly arranged grains involved in grinding. The aim of this research is to develop models for the statistical effects of multiple cutting grains. The key issue for the models is to derive the grain depth-of-cut for a distribution of cutting grains and specified operating conditions. In order to simplify the development work, plunge grinding (no cross feed) is being used for the prototype geometry.

Two different approaches have been taken to predict the grain depth-of-cut for given grinding parameters (grain size, shape, distribution, wheel radius, runout error, etc.). The first is an analytic model for steady-state grinding. The analytic model is based on the key assumption that there is no correlation between the surface profile generated and the profile of the active cutting grains on the grinding wheel. The workpiece surface position relative to the grinding wheel is then given only by its mean value and the surface profile itself is ignored. In the second approach, a computer simulation is used to generate the actual surface profile for steady-state grinding. The simulation takes correlation effects between the surface generated and the profile of the active cutting grains into account. The advantage of the analytic approach is its simplicity, the ease of application to a wide range of initial assumptions and the ability to provide closed form solutions. The computer simulation can predict representative surface features and can produce solutions for conditions which have surface correlation effects.

By combining this effort with previous work on chip geometry [1, 2] and microcutting tests for single cutting grains [3], a process model can be developed for grinding brittle materials. The intent is to either define the limits for eliminating subsurface damage (ductile-regime grinding) or to provide a rational basis with which to optimize the grinding parameters for control of subsurface damage in given grinding operations. The work summarized here represents results from the initial development of the analytic and computer simulation models.

ANALYTIC MODEL

A random distribution of spherical grains with radius R is assumed and the grinding wheel is assumed to be perfectly true (no runout error). The model is easily extended to arbitrary grain shapes and distributions as well as to runout error effects. The action of "average" cutting grains is shown in Fig. 1. The z axis projects downward from the wheel binder into the workpiece surface as shown. Since the actual surface profile is ignored, the position of the workpiece surface is given by its mean position $z_{surface}$. A_g is the projected area of the cut in the cutting direction and D_g is the mean diameter of the grain as measured in the workpiece surface. The maximum grain depth-of-cut relative to the workpiece surface is given by $d_m = z_m - z_{surface}$.

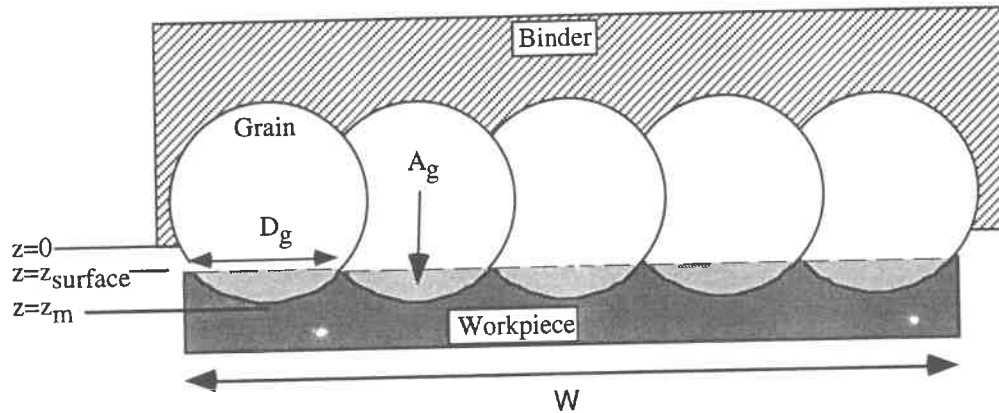


Figure 1 Schematic of plunge grinding.

The entire width W of the surface can be removed several times during a wheel revolution. The length of the grinding wheel that contains sufficient grains to remove one width W of the surface is called the characteristic length L and is a function of the surface position. The minimum number of cutting grains required to span one width of the surface, N_c , is given by W/D_g . The minimum length for L would contain only N_c grains. However a random distribution of grains is unlikely to be arranged so that the grains exactly span the width with no overlap, so a number of grains larger than N_c will be required to remove the width and L will be larger than the minimum value. The actual number of grains contained in a characteristic length, N_{tot} , is related to the number density ρ_{eff} of the cutting grains at $z = z_{surface}$. The total length of contact L_o of the grinding wheel is a function of the wheel circumference and the runout error. Assuming no runout error, L_o is just the circumference of the wheel ("continuous" grinding).

Under steady-state conditions for plunge grinding, the cross sectional area removed per wheel revolution by the cutting grains must be fW where f is the infeed per rev and W is the width of the cut. The cross section removed must then be equal to the number of times the surface width is removed, L_o/L , multiplied by the grain cutting area $N_c A_g$. In essence, the model treats continuous infeed of the grinding wheel as a succession of incremental motions where the duration for each increment is the characteristic length L . The relevant equations are

$$N_c = W/D_g$$

$$fW = \frac{L_o}{L} N_c A_g$$

$$N_{tot} = L W \rho_{eff}$$

$$f = L_o \rho_{eff} A_g \frac{N_c}{N_{tot}}$$

In order to derive the relationship between d_m and f , i.e., the grain depth-of-cut relation, explicit expressions are needed for A_g and D_g in terms of d_m . Let $g(z)$ be the grain distribution function defined such that $g(z)dz$ is the total number of grains between z and $z+dz$. $A(z, z_{surface})$ and $D(z, z_{surface})$ are geometrical functions giving the projected cutting area and grain diameter in the workpiece surface at given z , respectively. The relations for A_g and D_g then follow as

$$A_g = \frac{\int_{z_{surface}}^{z_m} A(z, z_{surface}) g(z) dz}{\int_{z_{surface}}^{z_m} g(z) dz} \quad D_g = \frac{\int_{z_{surface}}^{z_m} D(z, z_{surface}) g(z) dz}{\int_{z_{surface}}^{z_m} g(z) dz}$$

These can be solved to obtain the desired relation between f and d_m . The solution gives the mean workpiece surface position $z = z_{surface}$, but not the actual form of the surface profile. The detailed expressions are rather lengthy and will not be reproduced here.

There is a limit for the domain in which the analytic model is valid. Physical requirements demand that a sufficient number of cutting grains exist to "span" the width W of the cross section, otherwise uncut gaps will be left in the surface. If the contact length L_o is less than the characteristic length L , there are too few grains to remove a surface width and the model is no longer valid. The lower limit for the validity of the analytic model is established when the contact length $L = L_o$ and L contains exactly $N_c = W/D_g$ active cutting grains. One expects that there will be some overlap of the cutting grains, hence $N_{tot} > N_c$ and the characteristic length L is greater than the minimum value. Defining a characteristic length that takes the overlap of random grains into account is a complex statistical function, the analysis of which is currently underway. All of the results presented were obtained for the lower limit of the analytic model, where L contains exactly N_c grains.

COMPUTER SIMULATION MODEL

The computer simulation model was developed by Miller [1]. This model has been adapted and modified for the work reported here. Basically, the simulation method unwraps the grinding wheel and treats the operation like an inclined cutting plane, the slope of which is determined by the infeed. A random distribution of spherical grains with radius R is again assumed. An initial reference surface profile is established by making one revolution of the grinding wheel (at the given infeed) with the workpiece surface initially set at the binder surface. One additional wheel revolution then determines the steady-state grain depth-of-cut d_m , as well as parameters such as A_g and D_g . In order to obtain proper statistical averaging, a number of identical runs must be made using different random seeds for the generation of the grain distributions.

Fig. 2 shows results obtained from the computer simulation and the analytic model for the same set of input parameters. As was noted by Miller [1], the computer simulation results show two regimes. At lower infeeds, the relation is $d_m = f$, which is the simulation data with the higher slope in Fig. 2. Above a transition infeed, the fitted relation becomes $d_m = \text{constant } f^{0.4}$, which is the simulation data with the lower slope. There is a transition between these two regimes, the transition infeed being identified by the intersection of linear fits to the data points for the two regimes in the log plot. The limit line shown for the analytic model marks the limit of the validity of the model, which is the case where exactly one characteristic length L has cut the surface during one revolution of the grinding wheel ($L = L_o$). The analytic model is physically meaningful only for infeeds that lie to the right of this limit line. As noted earlier, the solution for the analytic model is equivalent to the section width W being removed L_o/L times, i.e., the continuous infeed is treated as a sequence of incremental infeeds. The analytical model gives a closed-form solution for the d_m vs. f relation. The solution is rather cumbersome for spherical grains, but it can be closely approximated by the relationship $d_m = \text{constant } f^{0.4}$, which is the same form that is obtained from a fit to the computer simulation results. The analytic model solution also includes all of the grinding parameters (grain size, concentration, etc.) in the proportionality constant.

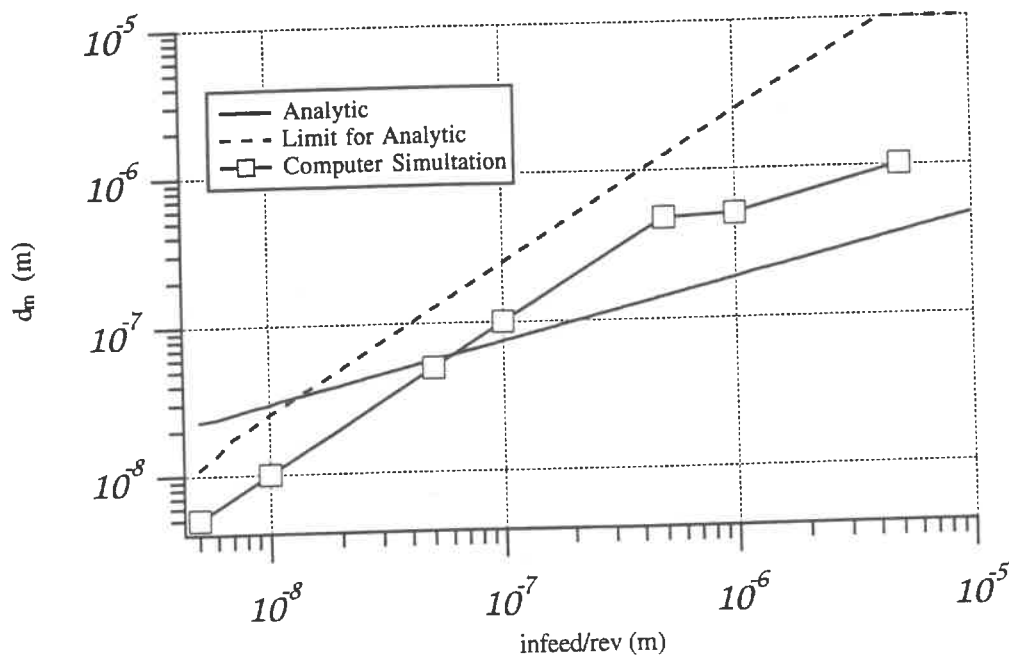


Figure 2 Model results for the maximum grain depth-of-cut vs infeed.

At lower infeeds, the simulation results give $d_m = f$. This is a result of the correlation between the surface profile and the active cutting grains on the grinding wheel. If the surface profile maintains a one-to-one correspondence with the cutting grains there is no grain interaction such that following grains cut into grooves formed by lead grains. Hence, each grain cuts a depth equal to the infeed and $d_m = f$. At higher infeeds, the relationship changes and $d_m < f$ due to the action of multiple cutting grains and grain interactions. The analytic model and the computer simulation results can be described by similar power-law relations for the higher infeed regime, although there is a relative shift upward (Fig. 2) for the simulation results due to the larger proportionality constant in the power law relation. Apart from this scaling factor, it appears that the surface correlation effects can be ignored at higher infeeds where grain interaction occurs.

The actual profiles cut into the workpiece surface can be tracked from the computer simulation outputs. For the low infeed case, there is almost a perfect correlation between individual cutting grains and the peaks and valleys in the surface profile. This implies the relation $d_m = f$ as noted above. In contrast, at higher infeeds there is a large amount of grain interaction and the surface profiles generated at various stages during one revolution of the grinding wheel have markedly different peak-to-valley features. There is no longer a one-to-one correspondence between individual cutting grains and grinding grooves; the spatial region bounding each groove in the final peak-to-valley profile has been previously cut by many grains and only the last grain leaves a signature of the groove.

While the analytic model does not apply to the lower infeed regime where surface correlation effects dictate the grain depth of cut relation $d_m = f$, it is noteworthy that the limit line for the analytic model falls rather close to the $d_m = f$ line for the simulation results (Fig 2). The limit line for the analytic model is based on the condition that the characteristic length L is just equal to the contact length L_o . Since by definition the characteristic length is the minimum length for

which the active grinding grains can remove a cross section width W in one revolution of the wheel, it seems reasonable that this limit should also be related to the onset of surface correlation effects. That is, below this limit the surface profile must maintain a strong correlation with the active grinding grains in order for the cross section to be removed. It also follows that in the low infeed regime, where $L < L_o$, if one starts with a perfectly flat surface there will be a transient period such that many revolutions of the wheel (infeeds) are required before the steady-state surface profile is generated. This is confirmed in the computer simulation results where it was observed that the peak-to-valley roughness can be much larger than the infeed.

The results presented here using the analytical model were obtained by setting the characteristic length L to its minimum value. The minimum L corresponds to the length needed for exactly $N_c = W/D_g$ cutting grains, the minimum number needed to span the width W in Fig. 1. For a random distribution of grains, the value of L must be larger than the minimum because there will be overlap of the grains. The statistical problem is equivalent to determining the average number of randomly tossed pennies of diameter D_g needed to just "cover" a line of length W . Preliminary estimates show that this number is approximately 4-6 times the minimum number. If this estimate is incorporated into the analytic model, the line for the analytic results in Fig. 2 shifts upward such that the analytic model and computer simulation results for the higher infeed regime are in good agreement. Furthermore, the limit line for the analytic model shifts to the right so that it gives a better estimate of the onset of surface correlation effects.

CONCLUSIONS

Preliminary results have been presented for modeling the grain depth-of-cut in plunge grinding. The key relationship derived is the maximum grain depth-of-cut d_m vs. the infeed per rev f . At low infeeds, there is a one-to-one correspondence between cutting grains and grinding grooves in the surface and the relationship $d_m = f$ is obtained using a computer simulation model that accounts for correlation between the surface profile and the profile of the active cutting grains on the wheel. At higher infeeds, $d_m < f$ because of the increased number of cutting grains and interactions between grains. These interactions reduce the correlation effects and change the form of the d_m vs. f relationship. A power-law relation $d_m = \text{constant } f^{0.4}$ is obtained for the higher infeed regime. The same form of power-law relation is also obtained using an analytic model which does not account for correlations between the surface profile and the active cutting grains on the grinding wheel. The analytic and simulation models differ only by a scaling factor. A lower bound limit for the applicability of analytic model gives a good estimate of the point at which surface correlation effects become significant. Current work is in progress to further refine and extend the models, and to develop experimental techniques to verify the theoretical predictions.

REFERENCES

- [1] Miller, M.H., "A Model for the Grinding of Brittle Materials", *PhD Dissertation*, N.C. State University, 1994, pp. 134-142.
- [2] Storz, G.E., "Theoretical Modeling of Precision Contour Grinding Operations", *MS Thesis*, N.C. State University, 1994.
- [3] Tanikella, B.V. and Scattergood, R.O., "Fracture Damage in Borosilicate Glass During Microcutting Tests", *Scripta Materialia*, Vol. 34, No. 2, pp. 207-213, 1996.

High-Efficiency Mirror Finishing Characteristics of Ceramics with Extremely Fine Grit Diamond Wheels

Xu-Jin Wang*, Katsuo Syoji** and Tsunemoto Kuriyagawa**

* SpeedFam Co., Ltd., No. 2647 Hayakawa, Ayase-city, Kanagawa-pref., 252, Japan

** Tohoku University, Aramaki-Aoba, Aoba-ku, Sendai, 980-77, Japan

1. INTRODUCTION

In order to solve the problems of the loose abrasive method, such as low finishing efficiency, etc., a new vertical type ultra-precision plane honing machine using an extremely fine grit diamond wheel was developed by the authors for mirror finishing of hard and brittle materials including ceramics as a bonded abrasive method. Using this machine, the high-efficiency mirror finishing of ceramics was carried out successfully¹⁾²⁾.

In this paper, the theoretical model and formulas of plane honing are presented along with the results of a few experiments for examining the high-efficiency mirror finishing characteristics of ceramics using extremely fine grit diamond wheels.

2. EXPERIMENTAL METHODS

Fig. 1 shows the main structure of the ultra-precision plane honing machine used in the experiments. In order to remove the effect of vibration, an air bearing was employed on the diamond wheel's spindle, and a hydrostatic bearing was also used on the workpieces' spindle in the machine. Moreover, for high-efficiency mirror finishing, the diamond

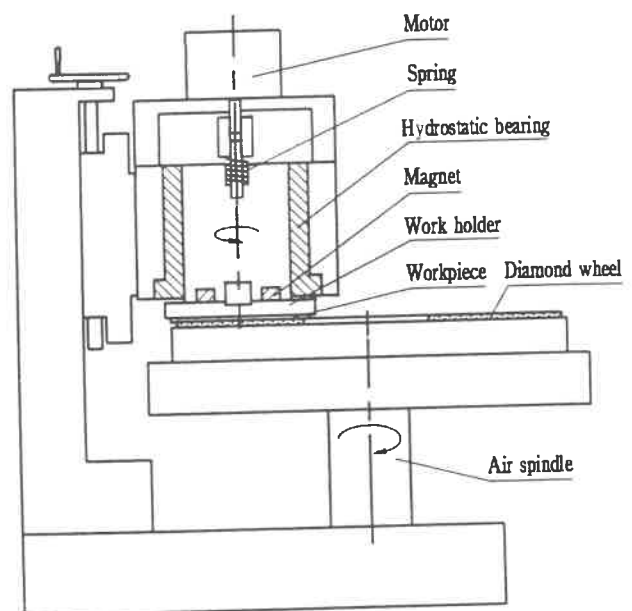


Fig. 1 Ultra-precision plane honing machine

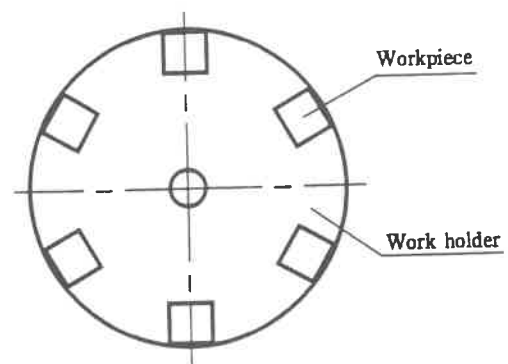


Fig. 2 Workpieces and work holder

wheel's spindle is a high-speed spindle which can be run up to 3000 rpm.

Fig. 2 shows the workpieces and holder. Prior to mirror finishing which uses extremely fine grit diamond wheels, the workpieces were lapped with diamond slurry. The workpieces were attached to the holder with glue in a concentric. The work holder was installed to the workpieces' spindle by magnets. The load was pushed down on the workpieces by a spring in the mirror finishing process.

To prevent the generation of hydro-planing force by the coolant, the plane type diamond wheels with a spiral groove was used in the process. Fig. 3 shows the diamond wheel. The truing and dressing of diamond wheels was performed by a cup type GC grinding wheel which was installed on the work holder.

3. EXPERIMENTAL CONDITIONS

In the experiments, a #1500 resinoid-bonded extremely fine grit diamond wheel and a #1500 metal-bonded extremely fine grit diamond wheel were used. Truing and dressing of these diamond wheels was carried out with a cup type #1500 GC grinding wheel.

The materials of the ceramic workpieces are zirconia (ZrO_2) and silicon nitride (Si_3N_4). The workpieces' roughness by lapping was R_{max} 1.5~1.7 μm prior to mirror finishing.

The experimental conditions are described in Table 1.

4. THEORETICAL MODEL AND FORMULAS OF PLANE HONING

The single point cutting model of plane honing is shown in Fig. 4. The grain on the extremely fine grit diamond wheel is supposed as a ball

Table 1 Experimental conditions

Grinding wheels	
Diamond wheels (ϕ 205mm)	① SD1500B100 ② SD1500M100
Truing wheel (ϕ 90mm)	GC1500H
Workpieces	
Materials	① Zirconia (ZrO_2) ② Silicon nitride (Si_3N_4)
Size	15mm $\times$ 15mm
Numbers of batch finishing	6 pieces (Work holder: ϕ 100mm)
Pre-machining method	Lapping (R_{max} 1.5~1.7 μm)
Finishing conditions	
Load	30 N
Workpiece speed	100 rpm
Diamond wheel speed	1000 rpm
Finishing time	5 min
Coolant	Water (15ml/min)
Truing and dressing conditions	
Load	30 N
GC wheel speed	300 rpm
Diamond wheel speed	300 rpm
Truing time	1 min

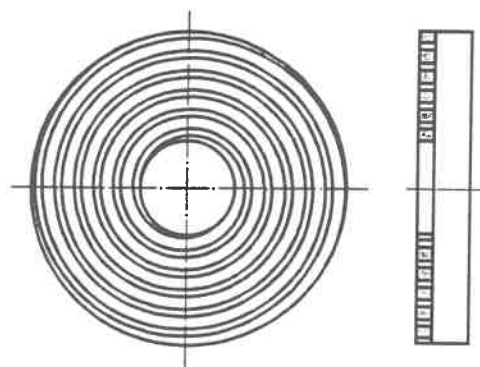


Fig. 3 Diamond wheel with spiral groove

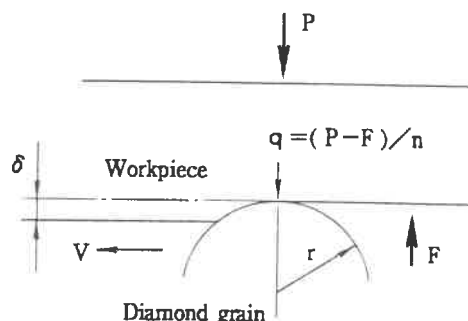


Fig. 4 Single point cutting model

with radius r in the model. The symbols in fig. 4 are represented as following:

- P: Finishing load
- F: Hydro-planing force by coolant
- n: Numbers of effective cutting grain in contact finishing area between workpieces and diamond wheel
- q: Cutting force of single grain
- v: Cutting speed
- δ : Cutting depth of single grain

The cutting depth of single grain δ in Fig. 4 can be calculated using Equation (1). H_v is the micro-hardness in finishing surface of the workpiece in Equation (1).

$$\delta = \frac{0.3(P - F)}{n r H_v} \quad (1)$$

The finishing rate h_w (the workpiece's removal depth per finishing time) can be calculated using Equation (2). A is the area of workpiece's finishing surface in Equation (2).

$$h_w = \frac{0.3V}{rA} \left(\frac{P - F}{H_v} \right)^{\frac{3}{2}} n^{-\frac{1}{2}} \quad (2)$$

5. EXPERIMENTAL RESULTS AND DISCUSSION

Fig. 5 shows the changes in the stock removal depth and surface roughness R_{max} according to the length of finishing time. during the finishing of zirconia (ZrO_2) using both a resinoid-bonded diamond wheel (SD1500B100) and a metal-bonded diamond wheel (SD1500M100).

Fig. 6 shows the changes in the stock removal depth and surface roughness R_{max} according to the length of finishing time during the finishing of silicon nitride (Si_3N_4) using both a resinoid-bonded diamond wheel (SD1500B100) and a metal-bonded diamond wheel (SD1500M100).

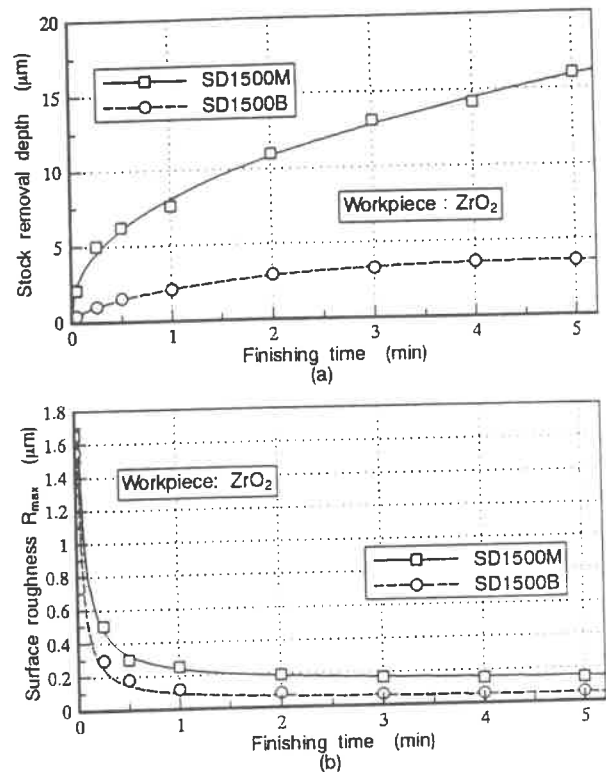


Fig. 5 Effects of diamond wheel's bonds (Workpiece: ZrO_2)

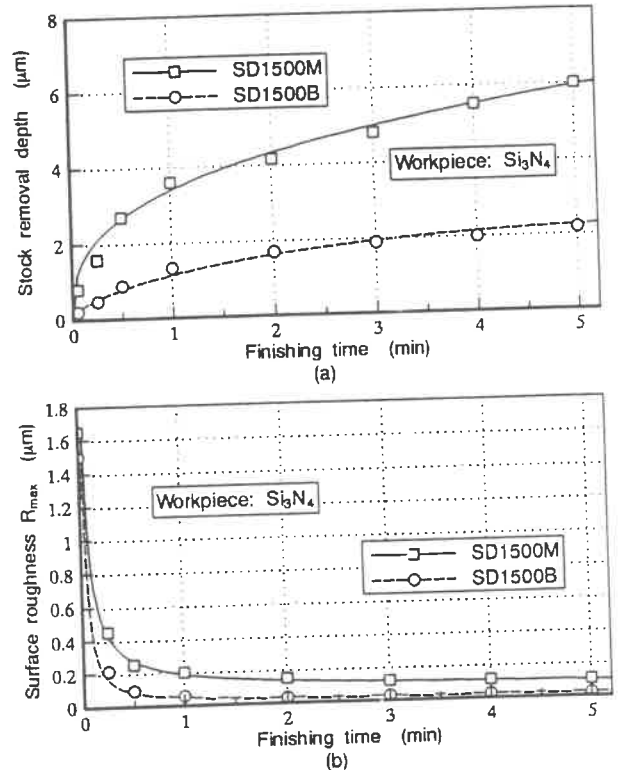


Fig. 6 Effects of diamond wheel's bonds (Workpiece: Si_3N_4)

We can understand that the finishing surface roughness using a resinoid-bonded diamond wheel is less than using a metal-bonded diamond wheel under the same mesh size and the same finishing conditions as stated in Fig. 5 and Fig. 6. The result is clear from Equation (1). Since the strength of the resinoid bond is less than that of the metal-bond, the numbers of effective cutting grain (n) on a resinoid-bonded diamond wheel is larger. As a result, the cutting depth of grains (δ) is less, so the surface roughness is less during the use of the resinoid-bonded diamond wheel.

However, Fig. 5 and Fig. 6 show that the finishing rate using a resinoid-bonded diamond wheel is less than that of using a metal-bonded diamond wheel. Using Equation (2) we can analyze the reasons. Since the strength of the resinoid-bond is less, the gap between the diamond wheel and the workpiece during the use of the resinoid-bond diamond wheel is less in the finishing process than during the use of the metal-bonded diamond wheel, so the hydro-planing force by coolant (F) is larger. Moreover, the numbers of effective cutting grain (n) on the resinoid-bonded diamond wheel is larger than that of the metal-bonded diamond wheel. For these reasons, the finishing rate (h_w) is less during the use of the resinoid-bonded diamond wheel.

Fig. 7 shows the comparative experiment's results of the final surface roughness $R_{\max}$ and average finishing rate during the finishing of zirconia (ZrO_2) and silicon nitride (Si_3N_4).

As shown by Equation (1) and Equation (2), the larger the hardness of the workpiece (H_v), the less the cutting depth of the grains (δ), and the less finishing rate (h_w). The results of the experiments shown in Fig. 7 agree with Equation

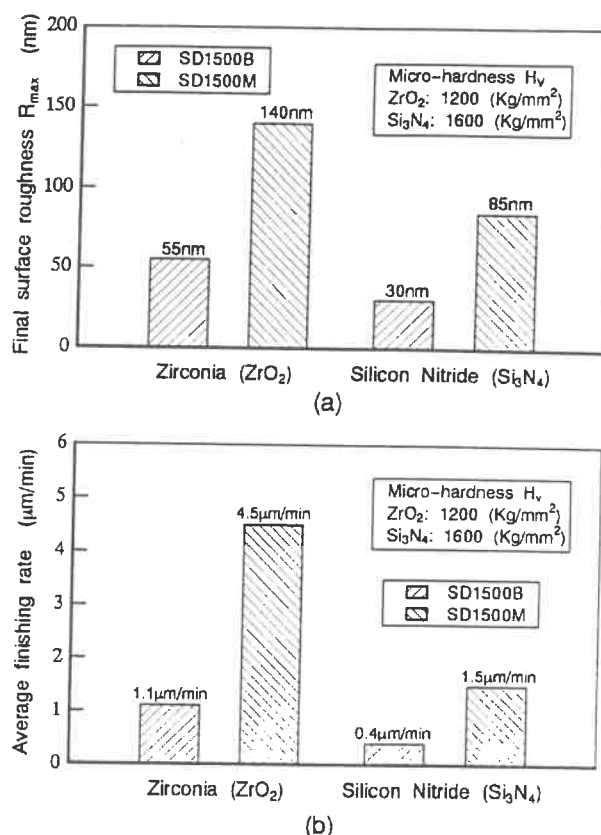


Fig. 7 Effects of workpiece's materials

(1) and Equation (2). Namely, the final surface roughness $R_{\max}$ and average finishing rate during finishing silicon nitride (Si_3N_4) are less than during finishing zirconia (ZrO_2).

REFERENCES

- 1) X. J. Wang, K. Syoji and T. Kuriyagawa: "High-Efficiency Mirror Finishing of Ceramics with Diamond Wheels", Proceedings of ASPE 10th Annual Meeting (1995.10.).
- 2) X. J. Wang, K. Syoji and T. Kuriyagawa: "High-Efficiency Mirror Finishing of Ceramics with Plane Honing", Proceedings of International Conference on Precision Engineering '95 (2nd ICMT) (1995.11.).

SHEAR-MODE GRINDING FORCE CRITERIA IN GRINDING SILICON WAFERS WITH CUP WHEELS

Y. Yoshida *, Pang Nan **, K. Imai***, S. Sekita^o, I. Iida^o,
N. Tsuda^{oo}, H. Hashimoto***

*Tokyo Diamond Tool(2-3-5 Nakane Meguro-ku Tokyo Japan), ** North eastern Univ.,

*** Kanagawa Institute of Technology(1030 Shimo-ogino, Atsugi-shi 243-02 Japan),

^oOkamoto-Shibayama Kikai, ^{oo}Tokai Univ.

1. Introduction

In recent years semiconductors have become more powerful and lower-priced, caused by an increase in the diameter of silicon wafers, with a tendency to grow from 12~16 inches in the near future. On the other hand, this has created a very serious problem in the manufacture of silicon wafers. In view of this fact, there are some trials to apply shear-mode (or ductile-mode) grinding for manufacturing of silicon wafers instead of lapping.

Previous papers ^{1) 2)} reported shear-mode grinding of various brittle materials using #3000 metal bond straight diamond wheels. In those experiments, it was shown that: (1) After a number of passes, the normal grinding force reached an approximate steady state; (2) during grinding, the wheel surface roughness and cutting performance were invariant; wheel truing and dressing were not carried out. From these experimental results, shear-mode grinding force criteria were defined as the steady, invariant grinding force which is a repeatable deterministic function of shear-mode grinding conditions including materials and grinding wheels. The criteria can usefully be applied to design high precision shear-mode grinding machines by determining their compliance.

We already tried to grind silicon wafers with #3000 metal bond diamond cup wheels. An approximate steady grinding force was found in grinding with serrated wheels with 1.0mm width. Normal grinding force over 25mm square test pieces was 10N at removal rate 10mm³/min. Grinding force ratio of tangential to normal was 0.15. Further, subsurface damage was investigated by photo-conductivity transient technique. The depth of subsurface damage in the shear-mode grinding at a removal rate of 10mm³/min was found to be 0.5-0.7 μ m³. But there is an indication that heavy metals of metal bond stick to the surface of silicon wafers.

This paper describes experiments with resin bond diamond cup wheels which do not include any metals as fillers, and actually manufacture silicon wafers on the production machine.

2. Experimental methods and results

The grinding machine is a reconstructed surface grinding machine (OKAMOTO PSG-52DX) with PI motorized block head shown in Fig.1. 25mm square silicon wafers (100 surface) were ground with cup wheels under conditions shown in Table.1. Normal and tangential grinding forces were measured using a three-axis quartz piezoelectric instrument.

Two types of resin bond serrated wheels (ϕ 100-1.0w) were used in the experiment. One of these had a general resin bond type-A, the other had a characteristic that is consistent with high concentration cores and low concentration foundation (type-B). In the first stage, large changes are caused at the resin bond wheel surface, with every experiment carried out after dressing with resin bond GC (#3000) cup truer under the same conditions.

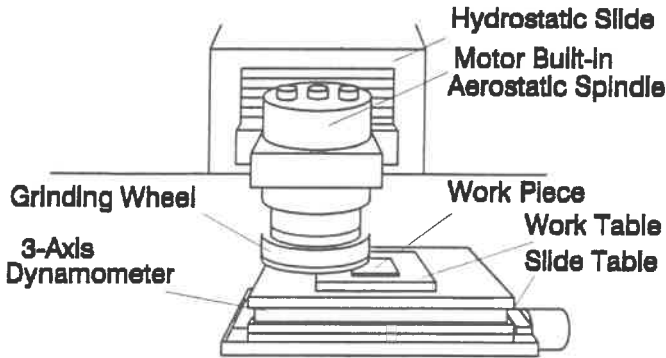


Fig.1 Schematic drawing of the grinding machine

Table.1 Grinding Condition	
Workplace	Sillcon(w25-l25-t2.5)
Diamond Wheel 100-1.0w (slit-48-3mm)	Resin bond wheel SD2000N100(Type-A) SD2000N100(Type-B)
Wheel Speed	750,1000,1500,2000 rpm
Table Speed	50, 100, 150 mm/min
Depth of Cut	1, 2, 4, 6 μ m/pass
Coolant	AL-1 (Yushiro Co.)

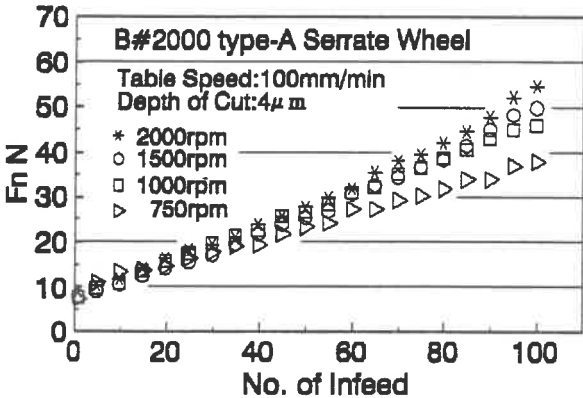


Fig.2 Normal Grinding Force of Type-A Wheel

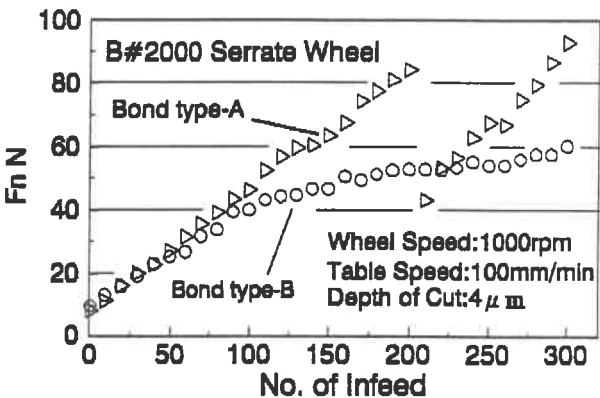


Fig.4 Comparson of Normal Grinding Force of Type-A Wheel with Type-B Wheel

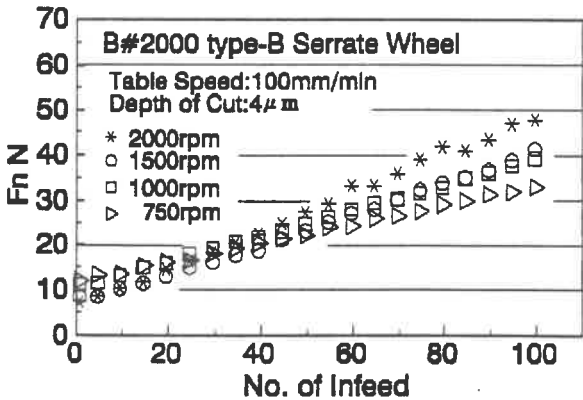


Fig.3 Normal Grinding force of Type-B Wheel

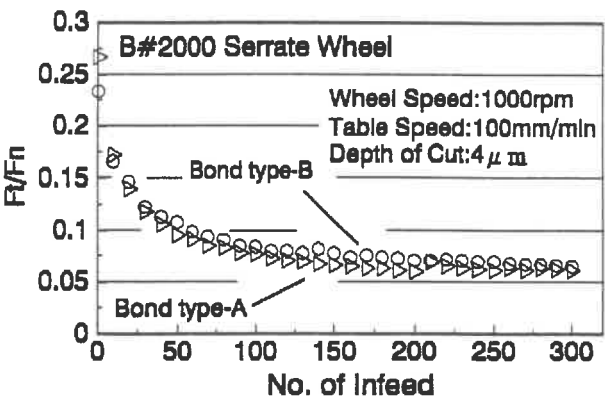


Fig.5 Comparson of Grinding Force Ratio (Ft/Fn) of Type-A Wheel with Type-B Wheel

Fig.2 and Fig.3 show change of normal grinding force in variable wheel speed. Normal and tangential grinding force increased with increasing of passes at each wheel speed. Increasing rate of normal grinding force was larger than tangential grinding force. Consequently, the grinding force ratio F_t/F_n came down with increasing passes. In variable table speed and depth of cut, normal grinding force was increasing with increase of passes and grinding force ratio F_t/F_n decreased gradually and approximated steady state of each grinding condition.

Next, we tried 300 passes and compared change of the grinding forces between two types resin bond diamond wheels. The normal grinding force increased quickly and created unstable change in the resin bond type-A. But it showed a stable change at the resin bond type-B.

3. Manufacturing of Silicon Wafers

We carried out manufacture of silicon wafers on the production machine, it is possible to grind from 3" to 8" wafers. This machine is a vertical spindle rotary table type.

Productivity and high accuracy on the part of the production machine is very important, accordingly the machine has two grinding stations, i.e. one is a rough grinding station with B#325 diamond wheel and another is a finish grinding station with B#2000 diamond wheel. And the machine has four vacuum chucks including stations for loading and unloading. To prevent micro cracks and dimples caused by contamination, the vacuum surface of the chuck is cleaned by de-ionized water and chuck cleaning devices. For example, the conditions of grinding for 8" silicon wafers is shown as follows: Diamond wheel: ϕ 250-3w-5x, SD#2000J75B, rotation of diamond wheel: 3000rpm, rotation of vacuum chuck: 50~20 μ m/min, feed rate: 25 μ m/min, stock removal of finish grinding: 20 μ m. It was possible to manufacture more than ten thousand 8" silicon wafers without any dressing. This fact shows the theory of the criteria has been applied to shear-mode grinding of 8" silicon wafers. Fig.6 shows a flatness of after grinding a 8" silicon wafer, realizing TTV=0.31 μ m.

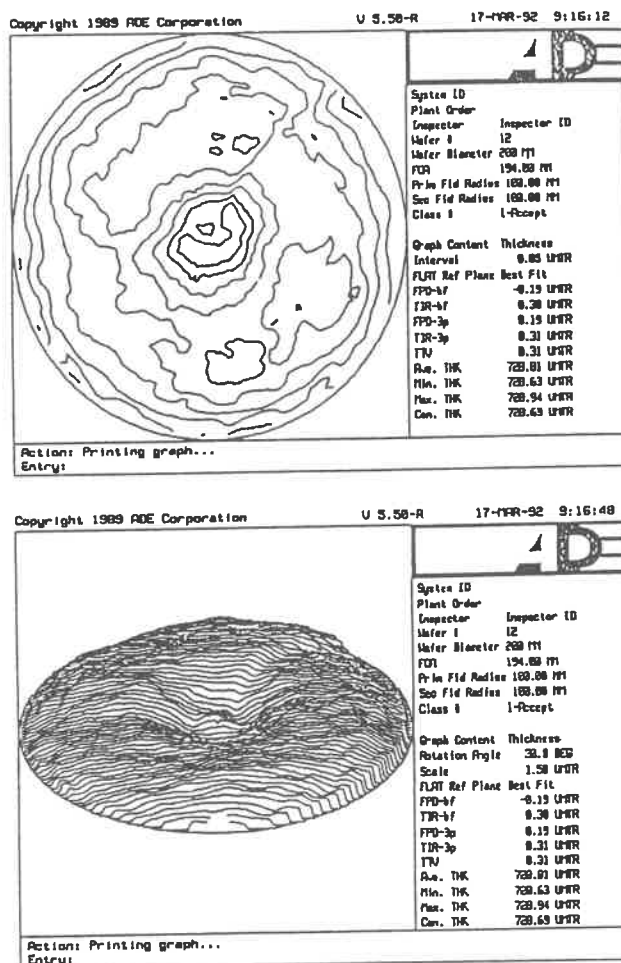


Fig.6 Flatness of ground 8" silicon wafer
TTV=0.31 μ m

An illustration of image measured by AFM is shown in Fig.7. It shows an experimental specimen, with grinding conditions as follows: Rotation of diamond wheel: 1000rpm, table speed: 100mm/min, and depth of cut: 2 μ m/pass. It shows some grinding streaks, but can't find any fractures or pits on the surface, it shows the finished surface was ground by shear-mode.

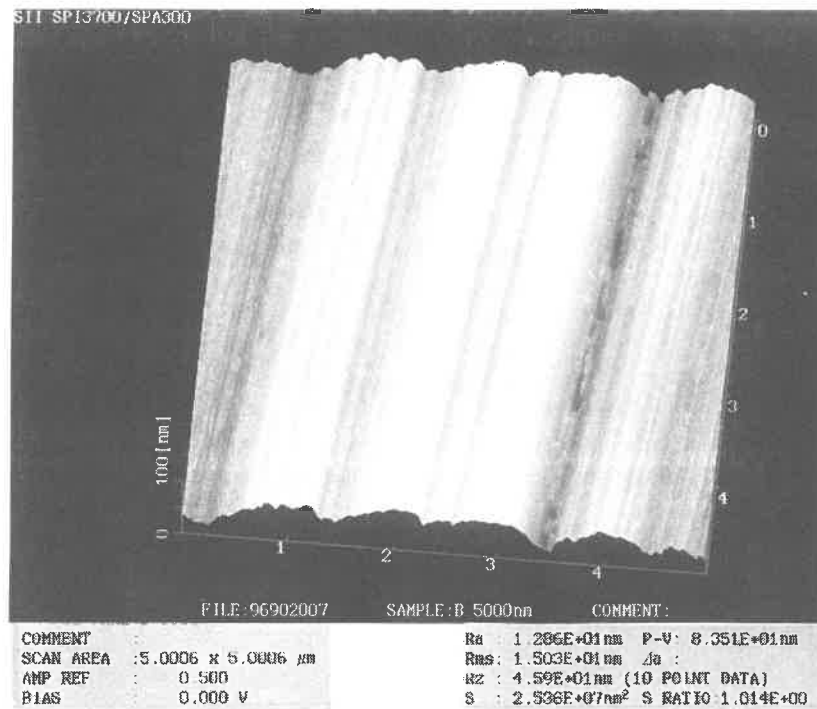


Fig.7 AFM image of ground with #2000 resin bond diamond wheel
Ra=12.8nm, P-V=83.5nm, Rms=15nm

4. Conclusions

1. An approximate steady grinding force like that of a metal bond wheel was not found with resin bond diamond wheels. But the grinding force ratio of tangential to normal was fixed under grinding conditions resulting in an increase of stock removal of close to 0.05-0.07.
2. Resin bond type-B with composite structure showed lower and stable normal grinding force in comparison with type-A.
3. It was found that more than ten thousand 8" silicon wafers can be ground without any dressing.

5. Reference

- 1) H.Hashimoto, J.Takeda, K.Imai, K.Blaedel: Shear-mode Grinding of Brittle Materials and Surface Characteristics, JSPE. 27,2(1993),95-99.
- 2) H.Hashimoto, K.Imai: Shear-mode Grinding Force Criteria in Grinding with Serrate Wheels, ASPE. (1995),171-174
- 3) Y.Ogita, H.Hashimoto, K.Imai, K.Yamamoto: The 42th Spring Meeting, 1995, The Japan

Cryogenic Grinding of Brittle Materials

Huadong YU, Shingo TERASHIMA, Noboru MORITA, Yoshitaro YOSHIDA

Chiba university, 1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba 263, Japan

1.Introduction

It is well known that the strength of brittle materials such as glass depends on the loading rate and loading duration. Usually, glass, when loaded at a rapid rate or forced to support a given load for short time, is relatively strong. In reverse, it is relatively weak. Indentation techniques, which have long been used as a means of measuring hardness of materials, are usually used to understand behaviors of material removal, deformation and fracture during machining brittle materials.^{1)-5).}

As the applied load gets higher, median crack, lateral crack, radial cracks appear around a residual impression formed by the indenter. Especially, it is thought that the cracks propagation depends on humidity of environment, and also depends on state of stress at the tip of the cracks.⁶⁾ When a crack occurred on the surface of the glass, a stress-dependent chemical reaction occurs between water vapor and the surface of glass. The stress corrosion promotes the crack propagation, and strongly depend on temperature of glass.⁷⁾ So we think that the stress corrosion will decrease with decreasing temperature of glass, and that it is possible to restrain the crack generation and propagation during machining brittle materials at low temperature. In order to get the fundamental information on the cryogenic grinding of hard and brittle materials, the diamond indentation tests and the single point diamond scratching experiments have been carried out at a room temperature and cryogenic temperature using liquid nitrogen respectively.

2.Experimental procedure

The tests of diamond indentation were carried out by using of micro-Vickers hardness tester at a room temperature and liquid nitrogen temperature

respectively. Material of test samples was soda-lime glass. The test specimen was kept on a special jig, which was full of liquid nitrogen and was placed on the table of the micro-Vickers hardness tester. The temperature of the indenter and the specimen were measured by thermocouple. When the specimen and the indenter had cooled down to liquid nitrogen temperature, the test indenter was loaded normally onto the surface of specimen with a dwell time of 30s. Indentations were for loads of 0.5 to 10 kgf. All observations were carried out at room temperature.

The single point diamond scratching tests were carried out at a room temperature and cryogenic temperature by using a grinding machine tool. In order to get continual depth of cutting, the test specimen was placed on the table of the machine tool with a small slope to the table. The scratching speed was determined by the feed of table of the machine tool.

3.Results and Discussion

3.1 Indentation

When the indenter indented into brittle materials such as glass, then above several kinds of crack usually occur. These cracks have also been observed in our experiments. Fig.1 shows dependence of radial crack length on indenting loads at room temperature and cryogenic temperature(77K) respectively. As shown in Fig.1, the crack length depends on indenting loads, and increases with increasing loads at both room temperature and cryogenic temperature. But the crack length occurred at cryogenic temperature was markedly smaller than one occurred at room temperature. Under an optical microscope, we observed that behaviors of deformation and cracking patterns at cryogenic temperature were different from those at a room temperature. At a room temperature, the boundary of

indentation (contact edges between the indenter and the surface of specimen) were very well defined, and some lateral cracks were found in subsurface by the reflections in addition to radial cracks. On the other hand, at cryogenic temperature, the contact edges were not clear, and the radial cracks decreased, but the ring crack occurred around the indented impression. Moreover the size of indentations were much smaller than those at room temperature. It seems that hardness of glass at cryogenic temperature is higher than that at room temperature.

The dependence of Vickers hardness of soda-lime glass on temperature is shown in figure 2. All tests were carried out during the specimen temperature rising up to room temperature, and all measurements were carried out at room temperature. It is found that the hardness of soda-lime glass markedly increases with decreasing specimen temperature. The hardness of soda-lime glass at low temperature (77K) was as high as about 2.5 times of that at room temperature. It is well known that strength of a material is directly proportional to its hardness. So we believe that strength of soda-lime glass can become higher considerably, when cooled to liquid nitrogen temperature. Moreover, it is found that there is a point of inflection at about 180K. Why for the point of inflection to exist can not be explained clearly. We think that the point of inflection may be due to a changing in such as chemical structure of soda-lime glass at about 180K. Naturally it is necessary to verify this view.

The relationship between the radial crack and the specimen temperature is shown in Fig.3. It is found that the radial crack decrease with decreasing specimen temperature, and that there is a point of inflection at about 180K as well. It is thought that the phenomenon about reduction of crack length may results from follows: 1) increasing in the fracture strength with temperature of specimen down; 2) reducing in water activity at low temperature.

Fig.4 is relationship between radial crack length and area of indentation in soda-lime glass. It is found that the radial crack length increased with the area increasing at both room temperature and liquid nitrogen

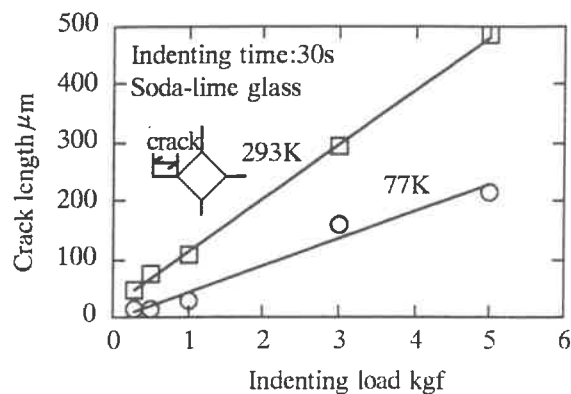


Fig.1 Dependence of radial crack length on indenting loads at room temperature and cryogenic temperature

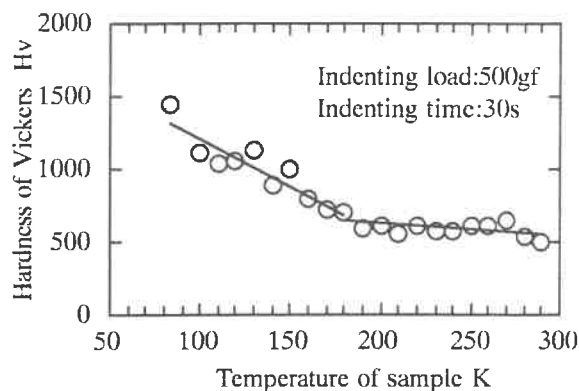


Fig.2 Dependence of hardness of soda-lime glass on temperature of specimen under room temperature

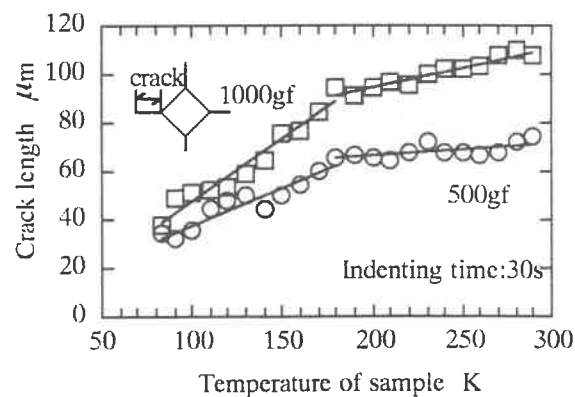


Fig.3 Relationship between radial crack length and specimen temperature under room temperature in soda-lime glass

temperature, but under just same area of indentation, the radial crack of the former is larger than the latter. In fact, because the indentation is deformed plastically and left on the surface of specimen by indenter, so the area of indentation amount to the volume of materials removed. It seems that the mechining at cryogenic temperature is favorable for restraining cracking ,in the case of just same materials removed.

3.2 Single point diamond grinding

Experiment of single point diamond scratching is very useful method to understand the mechanism of machining brittle materials, as we can directly get fundamental informations on behavior of the interaction between a abrasive and a specimen.

It is well known that the brittle materials deform plastically when the depth of cutting is very small. Generally, The depth of cutting, when cutting mode changed from brittle mode to ductile mode, is defined to critical depth of Brittle-Ductile Transition(denoted by d_c).

Fig.5 is dependence of critical cutting depth of BDT on the scratching speed. It is found that the critical cutting depth decreases with the scratching speed increasing, and depends on the specimen temperature. The d_c at the cryogenic temperature is larger than one at room temperature. This corresponds with the result of the indentation. Photograph of scratching grooves which corresponded to Fig.5 is shown in Fig.6. From Fig.6, we can find that region of ductile mode of groove scratched at cryogenic temperature, is larger than one at room temperature as well. And also we can find that region of ductile mode of groove depends on scratching speed and decreases with the scratching speed increasing. Moreover under an optical microscope, we could observe that a lot of lateral cracks and chipping appeared beside the grooves at both room temperature and cryogenic temperature, but region of the lateral cracks or chipping beside the grooves at the latter were smaller than at the former. It is thought that this phenomenon result from changing in strength of glass and water activity as the case of the indentation.

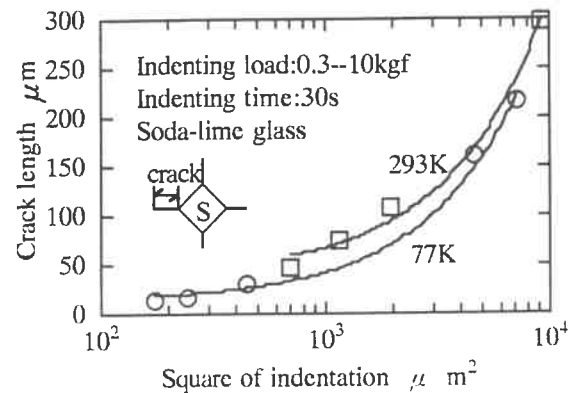


Fig.4 Relationship between crack length and area of indented impression

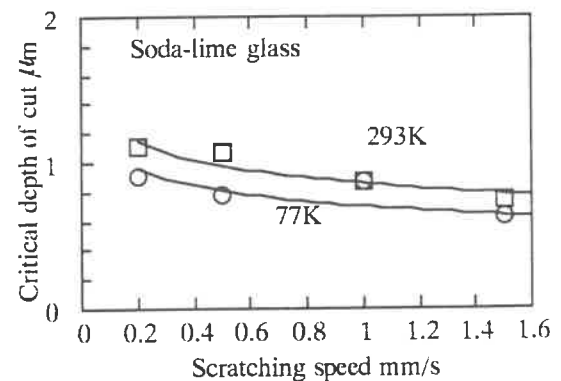


Fig.5 Relationship between critical depth of brittle-ductile transition and scratching speed

4. Conclusions

The experimental results of indentation tests indicate that when the temperature of specimen of soda-lime glass cooled down to cryogenic temperature (liquid nitrogen temperature), its Vickers hardness increases markedly, and that the radial crack due to indenter decreases. Moreover, the results of the single point scratching experiments indicate that at cryogenic temperature, the critical depth of cut without cracking is larger, and that the region of lateral cracks which occurred within subsurface of glass (or chipping on the surface of glass) is smaller than those at room temperature. So we believe that it is possible to restrain crack generation and propagation during grinding brittle materials such as glass at cryogenic temperature.

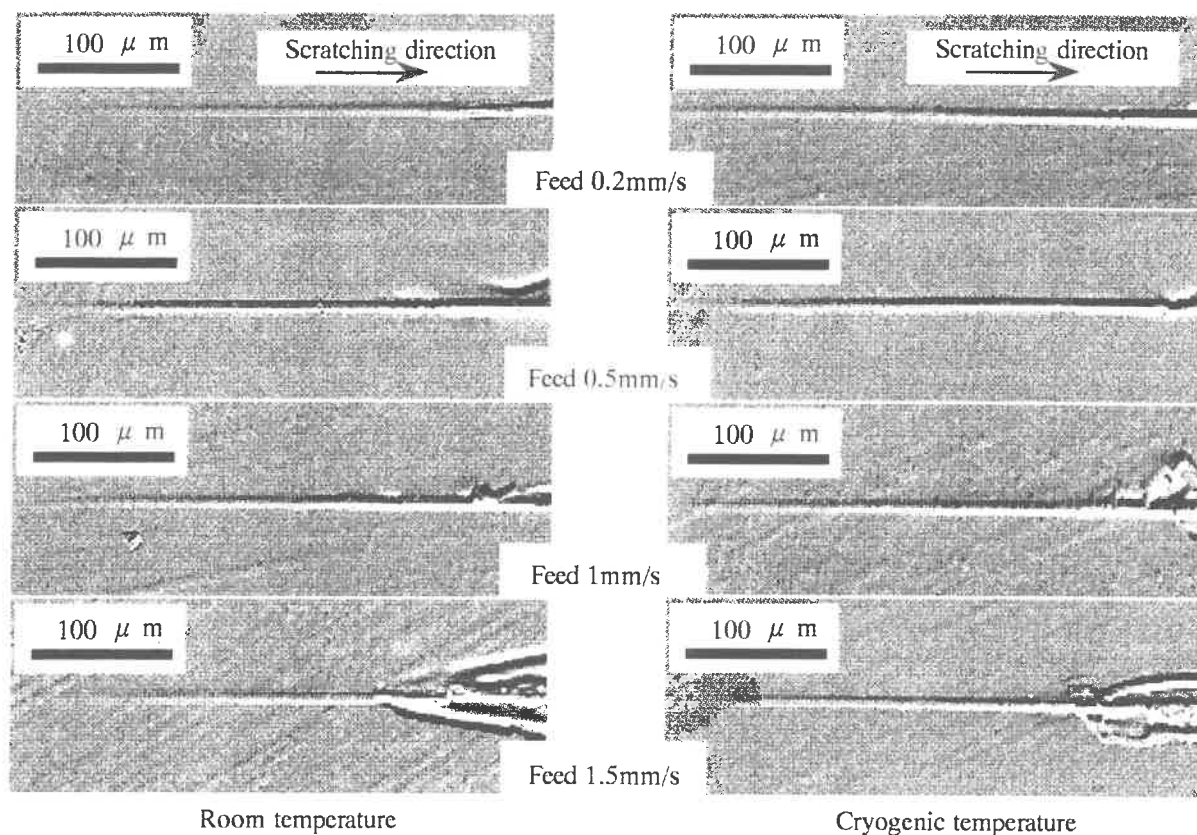


Fig.6 Optical photographs of scratched groove of soda-lime glass at room temperature and liquid nitrogen temperature respectively

5. Refrence

- 1) Maarten Buijs and Liesan A.A.G. Martens, Effect of Indentation Interaction on Cracking, *J. Am. Ceram. Soc.*, 75[10]2809-14(1992)
- 2) K. E. Puttick, M. A. Shahid and M. M. Hosseini, Size effect in abrasion of brittle materials, *J. Phys. D: Appl. Phys.*, Vol. 12, 1979
- 3) M. ARAGAKI, J. SHIBATA, Collection of Freatises in Autumn Meeting '1994 of THE JAPAN SOCIETY OF PRECISION ENGINEERING, 173--174
- 4) C. L. Chao, C. M. Fang, C. L. Chen, S. M. Tsay and S. S. Kuo, Investigation of the Micro-Mechanical Properties of Various Brittle Materials Using Micro-Indentation Technique, The Second International Abrasive Technology Conference, HsinChu, Taiwan, Nov. 1995
- 5) N. MORITA, Y. YOSHIDA, Effect of Temperature and Load Rate on Brittle-Ductile Transition of Silicon Single Crystal, ASPE'1994
- 6) S. M. WIEDERHORN, Influence of Water Vapor on Crack Propagation in Soda-Lime Glass, *J. Am. Ceram. Soc.*, 50[8]407-14(1967)
- 7) H. D. YU, N. MORITA, Y. YOSHIDA, Effect of Atmosphere on Crack Occurring and Extend during Indentation in Soda-Lime Glass, The Second International Abrasive Technology Conference, HsinChu, Taiwan, Nov. 1995

THE PLASTIC ZONE FORMATION IN ALUMINA ASSOCIATED WITH ULTRA PRECISION GRINDING

Irena Zarudi

Centre for Advanced Materials Technology, Department of Mechanical and Mechatronic Engineering, The University of Sydney, NSW 2006, Australia

1. Introduction

The ductile-regime grinding of alumina has been used in industry for several years, but still insufficient control of the process restricts its wide application. The problems arise from the fact that the mechanism of the ductile-regime grinding of alumina has not been fully revealed.

Recently it has been recognised [1] that a prominent plastic deformation is involved into the process and different slip and twin systems are initiated under the single-point ground groove. Furthermore, it has been concluded that the stress field in the vicinity of the cutting edge should be responsible for activation of these systems. Unfortunately, parameters of the stress field were not established and the mechanism formation of twin and slip systems hasn't been explored.

The present work is aimed at the computation of the stress field in the vicinity of the cutting edge as single-point grinding and computer simulation of the plastic zone formation during the process. Experimental data has been obtained as well and the theoretical results are compared with parameters of the real plastic zone formed while single point grinding of alumina in ductile mode.

2. Method

2.1 Theory. The model for calculation is based on the consideration of the plastic flow formation as a result of initiation and movement of dislocations and twins in the subsurface area. The crucial role in the calculation was the evaluation of the stress field (in particular, the shear stress distribution under the ground surface).

The shear stress evaluation. The shear stress component was computed in the plane perpendicular to the ground surface for 3-d elastic case and complex pyramid load distribution. The pyramid ramp distribution was worked out taking into account dimensions of a single-point groove (Fig.1). The base of the pyramid was approximated as an isosceles triangle with the side of the base equal to the width of the ground groove and the corresponding height equal to 0.5 of it's width. The volume of the pyramid was normalised by the applied load and then the height was evaluated. Once the load distribution had been determined, the analysis was conducted by applying normal and tangential loads separately and load was evaluated in terms of three harmonic functions [2]. Dydo and Busby [3] solution for constant load distribution was used and the final solution was found by method of superposition.

Computation of dislocations motion. The movement of dislocations was estimated in the plane perpendicular to the ground surface. Basal and prism plane dislocations were computed as they were identified in the subsurface regions of the ground alumina [1]. The shear stress component along possible direction of dislocations motion was treated as a driving force for dislocation movement and activation of the dislocation sources considered in the regions adjacent to the ground surface (up to 0.1 μm under the surface). Assume further the lattice to be ideal without any defects and the only force opposing the dislocation movement the friction force of the lattice. The force exerted on dislocation was determined in accordance with method [4] taking into consideration the interaction

of the adjacent dislocations. The final equilibrium position of dislocations was then evaluated. Effective resolved shear stresses (ERSS) are as following: 17 GPa for basal slip system, 5 GPa for prism plane slip system and 15 MPa for basal and rhombohedral twin systems.

2.2 Experiment. Two types of alumina with average diameter 1 μm and 25 μm were treated. Single point grinding experiments were conducted on a reciprocating sliding machine [1] to imitate the ductile regime mode of grinding. The diamond indenter with tip radius 100 μm and load 0.5 N was used. The microstructure of the subsurface was examined by transmission electron microscope [1]. The topography of the groove was investigated by means of atomic force microscope.

3. Results and Discussion

3.1 Shear stress. Results of the computation of the shear stress (Fig.2) revealed it's complex distribution with the absolute maximum in the middle of the groove and going sharply down along two symmetrical rims with deviation from the maximum point. The values of the shear stress along y axis follow the distribution of the shear stress along the groove width, distribution of the shear stress in z direction enables to establish stresses inside the material under the single-point ground groove. The contours corresponding to ERSS of the slip and twin systems initiated by the grinding process can be seen in Fig.3. Twin distribution was then approximated as an area with values of the shear stress exceeded 0.15 GPa (ERSS of twins). So twins can be expected under the whole width of the single-point groove and are spread to the depth of 19 μm .

3.2 Dislocations motion. Areas of possible dislocations motion (Fig.4) give the picture of the dislocation structure of the subsurface. Two zones can be easily explored. The first one with the width 3 μm and the depth 1.5 μm located in the middle of the groove where one can expect the possibility of the initiation of all 5 possible slip and twin systems. The second one covers the whole width of the groove (9 μm) and spreads to the depth of 4 μm . In the second area only 4 possible slip and twin systems can be expected. Van-Mises criterion can be satisfied only in the first area. The second one is dangerous from the point of view of the crack initiation.

3.3 Topography of the single-point ground grooves and structure of the plastic zones. The structure of the subsurface was explored before single-point grinding experiments to ensure that subsurface regions were dislocations free before tests. The depth of the grooves was measured to be 100 nm. The single-point ground grooves occurred to be smooth and fracture free with Rms roughness of the surface equal to 60 Å that is comparable with the values of the roughness of the polished surface. The structure of the plastic zones was as following. Three zones could be easily distinguished in the subsurface regions of 25 μm grained alumina (Fig.5a). The first region is characterised by extremely high density of dislocations. It is spread to the depth 1.5-2 μm under the surface. The second region is located under the first and covers the depth 5-6 μm under the surface. The density of dislocations in this region is much lower and single dislocations are resolved easily, some cracks are detected. Only twins were found in the third region, they spread to the whole length of the grains (up to 20 μm) beneath the surface. No dislocations were observed under the second region. Cracks were seen at the sites of twin-twin intersection. In the case of 1 μm (Fig.5b) grained alumina only the first zone with the high density of dislocations is distinguished in the subsurface region, it covers the first row of grains. No dislocation were detected in the second row of grains. It indicates that the initiation of dislocations occurs immediately under the ground surface and movement of dislocations is terminated by grains boundaries.

It follows that the first zone in the model corresponds to subsurface area with high dislocation

density, 5 possible slip and twin systems can be initiated there and this zone should be cracks free, the second zone corresponds to the zone in subsurface with low dislocation density, it is dangerous from the point of view of cracks initiation, twins can be observed up to depth of 20 μm according to theoretical and experimental results.

4. Conclusions

(1) The new model for plastic zone formation is worked out. It gives good correlation with the structure and parameters of the plastic zones formed as single-point grinding of alumina in ductile mode.

(2) Real ductile mode can be achieved while single-point grinding of 1 μm grained alumina. The structure of the plastic zone in this case is characterised by high density of dislocations and twins in the first row of grains. Von-Mises criterion is satisfied by initiation of the sufficient number of twin and slip systems.

(3) For 25 μm grained alumina single-point grinding with small depth of cut (100 nm) forms three regions in subsurface zone. The first of them is spread to the depth of 1 μm under the groove bottom and is micro cracking free. The second and the third spread accordingly to the depth of 5 and 19 μm under the groove bottom and can be dangerous from the point of view of micro cracks initiation.

Acknowledgments

The support of the ARC large grant scheme is appreciated. The author would like to thank her supervisor L. Zhang for guidance of this work.

References

1. I. Zarudi and L.C. Zhang, "The dislocation structure in alumina associated with ultraprecision grinding, Proceedings of the 10th ASPE Annual Meeting, 163(1995).
2. K. L. Jonsen, "Contact mechanics" (Cambridge University Press, New-York, 1985).
3. J. R. Dydo, H. R. Busby, "Complete 3-D elasticity solutions in the half-space for constant and linearly varying pressure loads", Proceedings of the international Conference on contact mechanics. 19-26 (1993).
4. C.M.Li, "Dislocation Modelling of Physical Systems", edited by C. S. Hartley, M. F. Ashby, R. Bullough and J. P. Hirth (Pergamon Press, Oxford, 1981).
5. T. G. Bifano, T. A. Dow, R. O. Scattergood, "Ductile-regime grinding: a new technology for machining brittle materials", Transactions of the ASME, **113**, 184(1991).

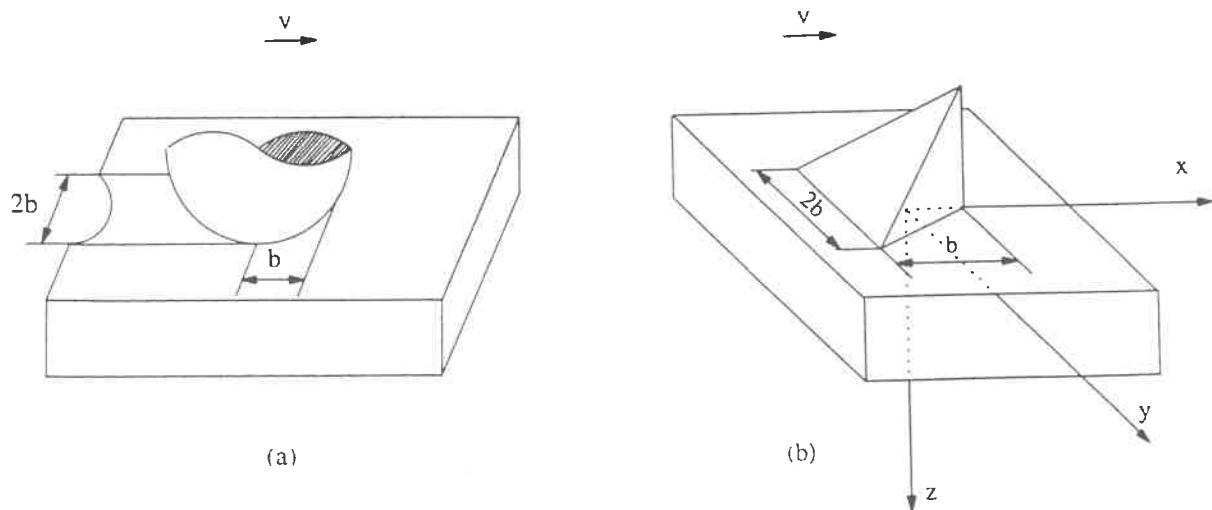


Fig.1 The model for shear stress calculation. (a) The single-point grinding. (b) The pyramid load distribution model.

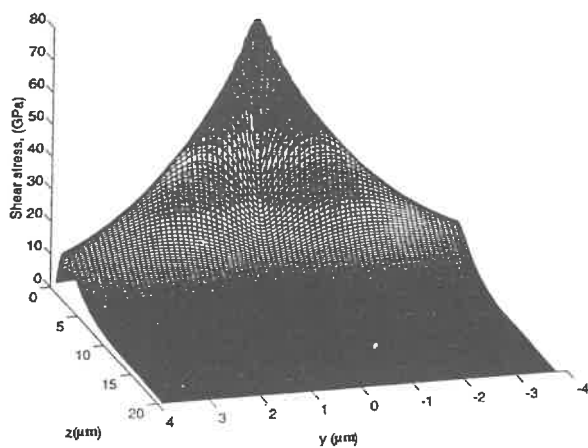


Fig.2 Shear stress distribution in yz plane.

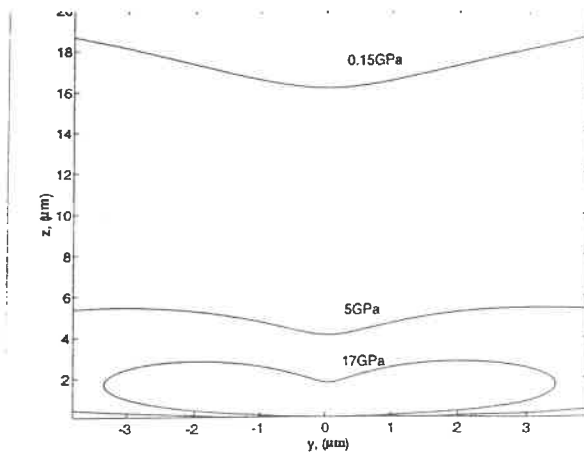


Fig.3 Contours of equal shear stress in yz plane.

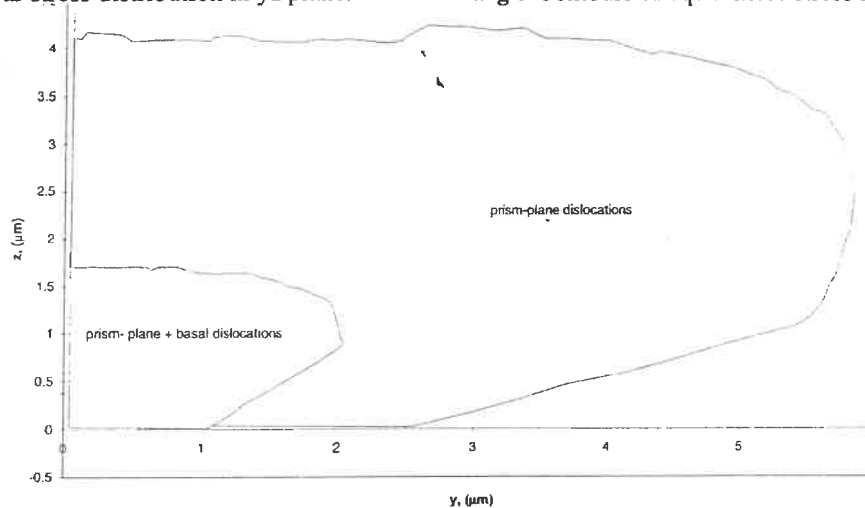


Fig.4 Areas of possible dislocations motion under the single-point ground groove.

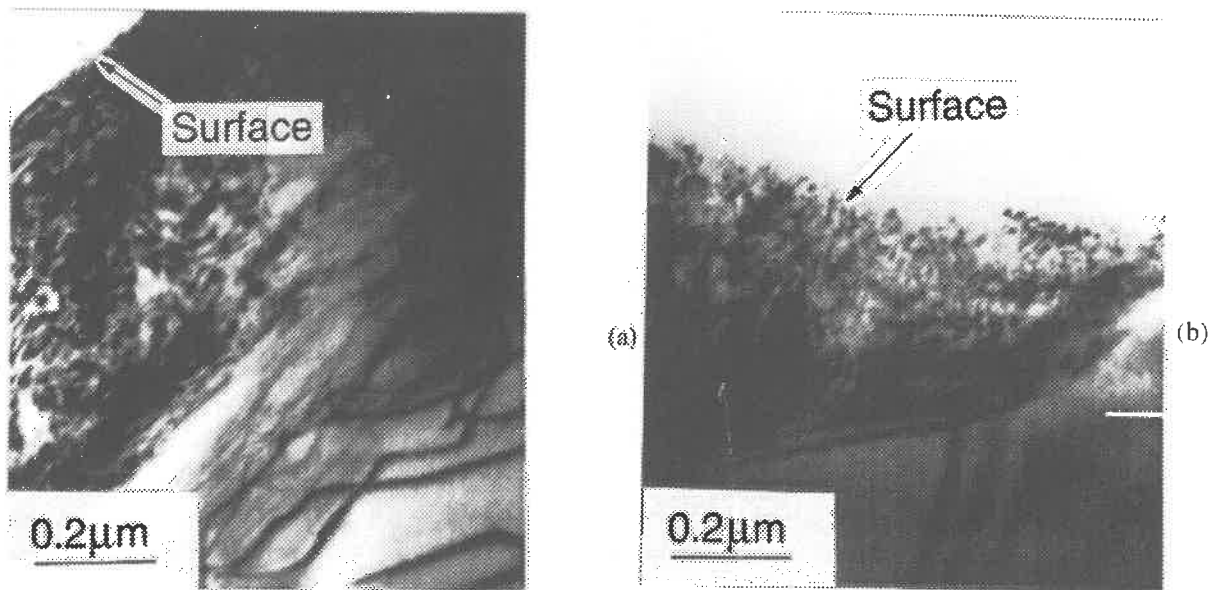


Fig.5 Subsurface damage in (a) 25 μm and (b) 1 μm alumina after single-point grinding tests.

Friction Characteristics of Linear Plain Bearing Guideway and Motion Controllability of Numerically Controlled Slide

A. Kanai, M. Miyashita, T. Hatai, M. Yoshida

Ashikaga Institute of Technology, 268 Omae, Asikaga, Tochigi 326, Japan

1. Introduction

Feed in NC machine tool is composed of summation of small stepwise feeds. Frictional load characteristics of plain bearing guideways for such stepwise feeds is expected to be different from it for continuous feeds. In the present paper the frictional load characteristics for stepwise feeds of 1 nm/step, 10 nm/step, 100 nm/step and 1 $\mu\text{m}/\text{step}$ at averaged feed rate of 0.05 $\mu\text{m}/\text{sec}$ to 5 $\mu\text{m}/\text{sec}$ in reciprocating slide feed is studied.

2. Test rig

Fig. 1 is the test rig. The slide is vertically supported by plain bearing guideways and horizontally by hydrostatic bearings. The slide positioning servo system is composed of force-operated hydraulic actuator and capacitive position sensor as feedback sensor. The slide is 450 mm $\times$ 300 mm in dimension and 150 kg in mass. The plain bearing guideways are two flats of 25 mm width $\times$ 450 mm length. Lubrication is applied with oilbath and ISO 22 oil is supplied as lubricant and power source. Piston of the hydraulic actuator is hydrostatically supported and pressure difference between the two cavities is perfectly

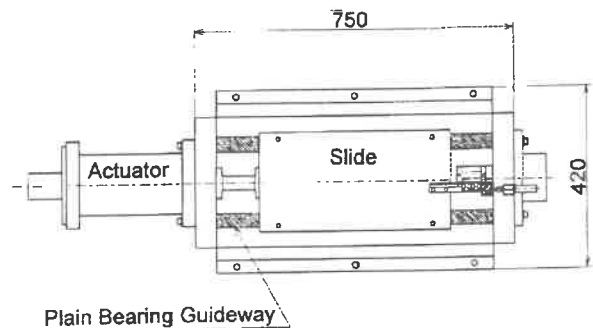


Fig. 1 Test rig

transmitted to the slide as driving force. The driving load is the inertia of slide and the friction, and the inertia is negligible compared with the friction except at the transient. And the frictional load of the plain bearing guideways is derived from the pressure difference between the two cavities.

3. Motion and frictional load characteristics for continuous feeds and stepwise feeds of 1 nm/step and 10 nm/step

Fig. 2 shows the test results of slide displacements and frictional loads in time for the reciprocating feed of 2 μm at the feed rate of 0.5 $\mu\text{m}/\text{sec}$. By reversing feed direction the frictional load increases with the displacement to the stationary state. The region in which the frictional load increases with displacement after reversing is defined as the transient one and the region in which the frictional load arrives at the stationary state is defined as the stationary one.

Fig. 3 is arranged for explanation of the transient and stationary regions in detail. They are the test results of the slide displacements and the frictional loads for the continuous feeds in the reciprocating feed range of 16 μm at the feed rate of 0.5 $\mu\text{m}/\text{sec}$. (a) shows the frictional characteristics after reversing feed direction from backward to forward and (b) does the one after reversing from forward to backward. Both varying process

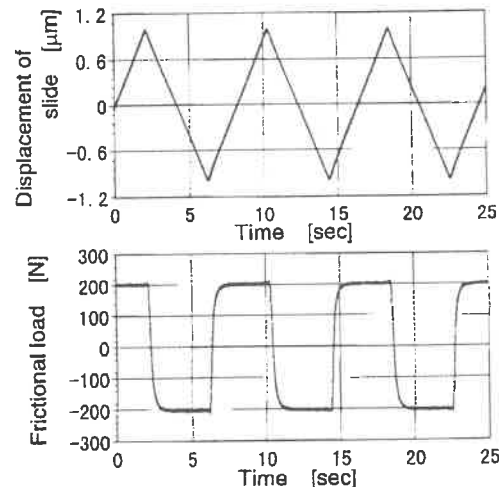
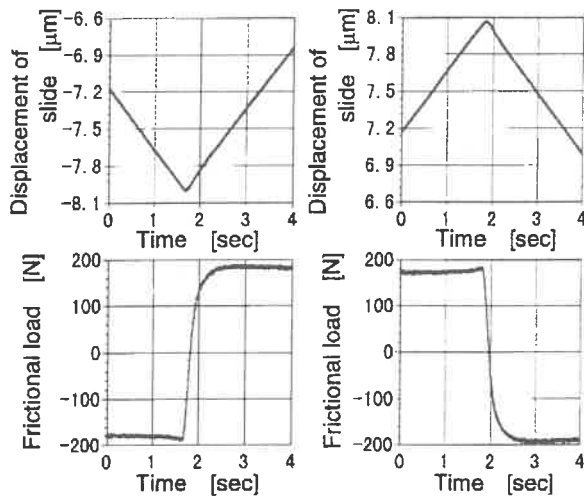


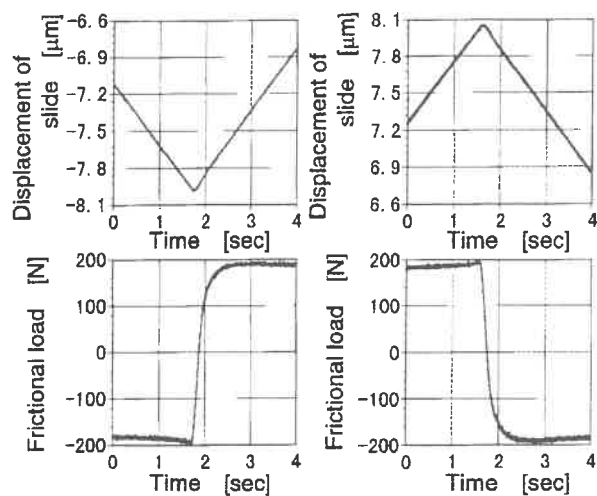
Fig. 2 Frictional Characteristic for Reciprocating Continuous Feed Motion (0.5 $\mu\text{m}/\text{sec}$)

of frictional load coincides very well. The 90 % of the slide displacement arriving at the stationary frictional load is assumed as a range of the transient region and it is about 0.2 μm .

Fig. 4 shows the test results of the slide displacement and the frictional load around reversing the feed direction for the stepwise feed of 10 nm/step at the feed rate of 0.5 $\mu\text{m}/\text{sec}$ in the reciprocating feed range of 16 μm . The relationship between the slide displacement and the frictional load obtained for 10 nm/step slide feed is also just same as the one in Fig. 3. (a) shows the one from backward to forward and (b) does from forward to backward. The transient region is about 0.2 μm . Test results in Fig. 3 and 4 coincide respectively very well and the stationary frictional load is both about 170 N. Here, the relationship is defined as "displacement vs. frictional load characteristic curve".

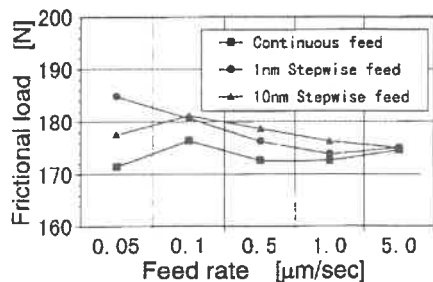


(a) Backward→Forward (b) Forward→Backward
Fig. 3 Transient Frictions at Reversing Point of Continuous Feed Motion (0.5 $\mu\text{m}/\text{sec}$)

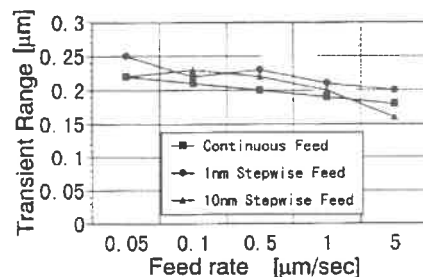


(a) Backward→Forward (b) Forward→Backward
Fig. 4 Frictional Characteristic for Reciprocating Stepwise Feed Motion (10 nm/step, 0.5 $\mu\text{m}/\text{sec}$)

Fig. 5 (a) shows the frictional loads in the stationary region for the continuous and stepwise feeds in the reciprocating slide feed at constant feed rate from 0.05 $\mu\text{m}/\text{sec}$ to 5 $\mu\text{m}/\text{sec}$. Fig. 5 (b) shows the ranges of the transient region. The feed rates are 0.05, 0.10, 0.50, 1.0 and 5.0 $\mu\text{m}/\text{sec}$ and the travel range is 16 μm . Difference of the frictional loads at the same feed rate between continuous and stepwise feeds is 8 % at most and the one at the different feed rates is 6 % in maximum. The frictional load tends a little bit to decrease with increase of the feed rate, and the range of transient region does similarly in the range of 0.25 μm to 0.16 μm .



(a) Friction in Steady State Domain



(b) Transient Range in Displacement

Fig. 5 Frictional Load in Steady State Domain and Transient Range in Displacement

4. Motion and frictional load characteristics for slide feed of 100 nm/step and 1 $\mu\text{m}/\text{step}$

Frictional load characteristics for stepwise slide feeds of 100 nm/step and 1 $\mu\text{m}/\text{step}$ in the reciprocating feed

for the cases of constant dynamic friction and velocity dependent friction were developed. The values of $\mu_{\text{static}}=0.25$ and $\mu_{\text{dynamic}}=0.35$ were used for the results of the simulated friction event in Figure 3.

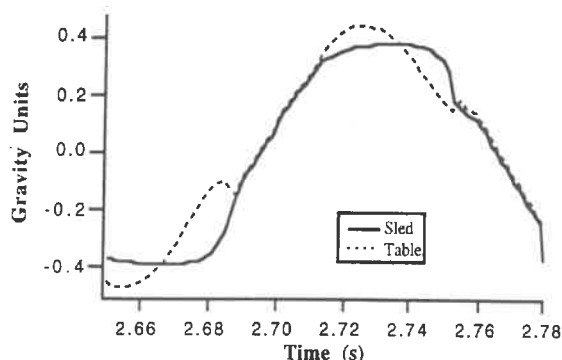


Figure 3: Simulated increasing velocity dependence.

This figure shows the acceleration measured during half of one cycle of motion where the sled (solid line) breaks away from the table (dotted line). Notice that near 2.69 seconds the curves become coincident indicating that the sled and table "stick" together as the acceleration (and therefore inertia) is reduced. As the acceleration increases again, the sled breaks away at about 2.71 seconds with a friction coefficient of approximately 0.25. This simulation uses a friction coefficient that increases with relative sliding speed ($C=5$).

Compare this trace with that in Figure 4 measured using a mating pair of steel specimens coated with a dry film lubricant (Hi-T-Lube) supplied by General Magnaplate. General Magnaplate suggests Hi-T-Lube-on-Magnaplate HMF™ for plain slideway bearings to reduce stiction.

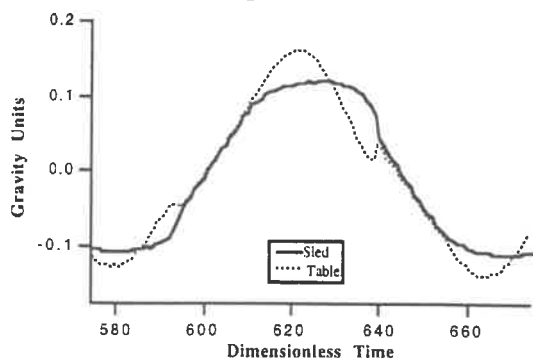


Figure 4: Measured Hi-T-Lube-on-Magnaplate HMF™ sled and table accelerations.

Figure 4 shows a friction event for this bearing counterface combination. The Magnaplate engineering data brochure claimed an elimination of stiction motion. If Figures 3 and 4 are compared, it can be seen that the simulation does model the system dynamics very well, and the Magnaplate bearing counterface does eliminate the effects of stiction. Also, Figure 3 demonstrates that the experimental results of Figure 4 can be duplicated by using an increasing velocity dependence for the dynamic friction.

Figure 5 shows the velocity dependence of the Magnaplate counterface. This curve was calculated from the acceleration data shown in Figure 4 by integrating the difference between the table and sled acceleration. The arrows on the curve indicate the integration path. The top part of the curve corresponds to the portion of Figure 4 where the table acceleration (dotted line) is increasing relative to the sled (solid line) just after breakaway.

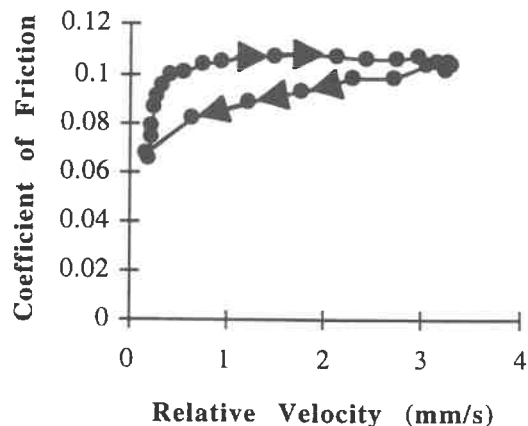


Figure 5: Range of dynamic friction.

This shows the dynamic friction increasing with respect to the relative velocity, much like in Figure 2. The maximum relative velocity between the sled and table occurs here at 3.5 mm/s, then the table decelerates relative to the sled. This corresponds to region of Figure 4 where the sled and table curve cross until they re-couple. The hysteresis in Figure 5 demonstrates that the frictional characteristics of an interface are dependent on whether the bodies are accelerating or decelerating relative to each other.

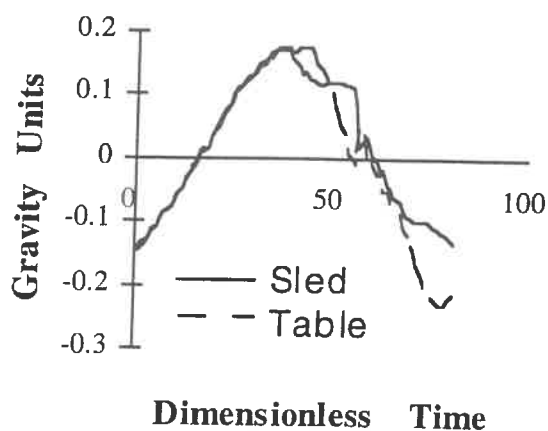


Figure 6: Vespel-on-Vespel

For comparison, the frictional properties of Dupont Vespel SP-211 were measured by sliding Vespel-on-Vespel. The friction event is shown in Figure 6. The sled acceleration curve has been overlaid on the table curve for comparison. The sled curve is a solid line and the dashed curve is the table. Initially, the curves are equal until the sled breaks free, at $\mu_{static} \cdot g$. The acceleration of the sled then drops to a virtually constant value $\mu_{dynamic} \cdot g$. A bearing material interface showing these characteristics can be described as a material with stiction. The values attained for this test yield $\mu_{static}=0.18$ and $\mu_{dynamic}=0.12$ which agree well with the values given in the Vespel Properties Bulletin which are $\mu_{static}=0.20$ and $\mu_{dynamic}=0.12$.

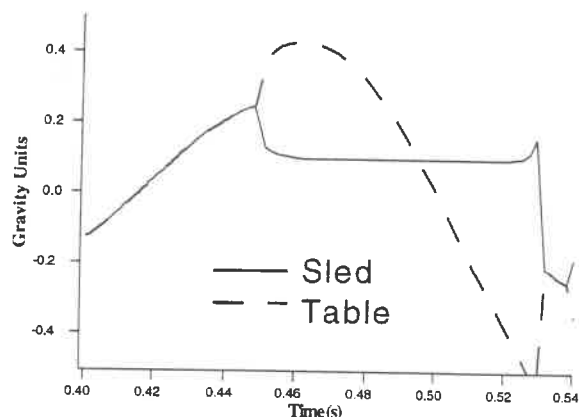


Figure 7: Simulated decreasing velocity dependence.

A calculated friction event with $C = 5$ is shown in Figure 7. The transition from breakaway friction to sliding friction is simulated as a decreasing function with relative velocity. This resembles the measured Vespel data of Figure 6 in that there is a distinct breakaway and a transition to

constant sliding. The perturbation in the sled acceleration just as it stops sliding relative to the table is due to the assumption in the simulated dynamics that the breakaway friction is the same as the termination (of sliding) friction. The experimental results attained here indicate that this may not always be true.

Friction measurements were performed on a variety of interfaces, including traditional bearing interfaces (PTFE, disulfide), Diamond-Like Carbon (DLC), and common materials (steel, glass). The PTFE bearings, which are known to reduce stiction, were shown to have an increasing velocity dependence but the disulfide coatings had stiction. The DLC materials showed velocity dependence that varied with deposition parameters and composition.

SUMMARY

The scope of this study has been to measure and to characterize the frictional behavior of candidate bearing coatings. A unique tribotester was developed and modeled so that the frictional behavior of bearings surfaces could be characterized. It is important to note that the Vibrational Tribotester developed in this research measures the transient frictional behavior during breakaway and eventual coupling (sticking) of the sled and table system. In that regard, it is different from the conventional, steady state friction tests, for example pin-on-disc. The system is designed to provide a unique characterization of breakaway friction behavior under conditions simulating machine tool bearing surfaces. The measured acceleration(s) can be integrated to yield a profile of the velocity dependence of the interface. However, the accelerometer output must match for each time step to be completely accurate. In that regard, the integral of the acceleration up to steady-state, where the table acceleration curve crosses the sled curve, is the most reliable for interfaces with decreasing velocity dependence.

Non-Isotropy of Young's Modulus of Elasticity in Very Thin Metal Sheets

P.V.Pistecky

Delft University of Technology
Laboratory for Micro Engineering
2628 CD Delft, The Netherlands
email: p.v.pistecky@wbmt.tudelft.nl

Introduction

High precision devices used in research, development or production often require a very accurate spring action for their proper function. In many cases, very delicate thin metal foil springs must be applied to establish the necessary spring function. Very often these springs possess rather complicated shapes, usually as a result of space restrictions.

In the past three years, several types of thin foil springs have been implemented in devices and instruments designed by the Laboratory of Micro Engineering of Delft University of Technology. The NASA Gamma Ray Collimator¹ and the Ultra Sensitive Force Transducer for Biomedical Experiments² are two examples extensively using this type of springs. The typical thickness of the metal foils used in our devices ranges from 20 μm to 100 μm .

To design and calculate the characteristics of these types of thin springs, finite element analysis is often the only adequate tool to obtain sufficient accuracy. However, comparisons between calculated results and accurate laboratory measurements on real springs showed repeatedly considerable differences. After careful investigations it became obvious that the non-isotropy of Young's modulus of elasticity of applied thin metal sheets is most likely the reason for this mismatch.

For some practical reasons most of the springs used in our devices are made from stainless steel thin cold rolled sheets. Due to the cold rolled process and the need for high allowable stress of the spring material, the sheets apparently possess different material structure in the rolled direction and perpendicular to this direction. The material properties depend obviously very strongly on the length, orientation and number of crystals throughout the structure. The tensile strength of the thin stainless steel sheets (AISI 301) reaches a value of 2200 MPa for the thickness of 15 μm .

Material properties

Instrument designs very often require an accurate prediction of the component's behavior before the actual realization of the instrument itself. This is due to the fact that many scientific instruments are expensive, single piece, unique prototypes. For this reason, it is almost impossible to check the instrument properties by building a pre-prototype version of that instrument.

During the last five years our research group gained substantial experience in using the finite element method for an accurate prediction of mechanical properties of instrument parts, especially with respect to their mechanical stiffness. It is obvious that accurate calculations are only possible if all material properties are known with sufficient accuracy. If a material possesses a non-isotropic distribution, for example, of Young's modulus of elasticity, then this distribution is of great interest.

Measurement of the non-isotropy of Young's modulus of elasticity

As mentioned previously, the cold rolled metal sheets possess different material structure and therefore also different Young's modulus values in the rolled direction and in the direction perpendicular to the rolled direction. In the first approximation, it was reasonable to determine the Young's modulus values in only these two main directions.

For all measurement experiments it was necessary to have some suitable specimens. In most cases they were shaped as long narrow strips with a length axis along the rolled direction or perpendicular to it. All specimens used for our experiment were made by our precision metal etching technique. With this

manufacturing procedure it is possible to obtain shapes in the order of tens of millimeters with an accuracy within a few micrometers. Due to this dimensional accuracy, the tolerance of the specimen dimensions does not play any substantial role in the possible measurement errors. The etching technique has an important additional advantage of a stress-free manufacture.

To investigate the modulus values, different measurement methods had been used in the past two years. Initially, the measurements started with the strain gauge method, but this method failed for most of the thin foils due to the relatively large stiffness of the strain gauges.

The next method used in our investigations was the measurement of deflection of a cantilever spring tip loaded only with its own mass. Even with very simple means, it was possible to obtain rather accurate results that could be compared with the method of natural frequency described later. A discussion about the measurement results will be presented in an other paragraph.

Coil set of the rolled sheets

By using the method of cantilever beam deflection, it became obvious that the slight sheet curvature in the rolled direction, due to the coil set, can cause measurement errors. If the sheet curvature is perpendicular to the beam length axis, the moment of inertia of the beam cross section can reasonably be larger than that of a flat sheet. Therefore some calculations were made to establish the possible influence of this phenomenon. The only way to avoid this influence is by choosing the beam width which is small enough. Figure 1 gives some results of the ratio of the moments of inertia of curved and flat sheets. As the radii of the coil set in all cases were larger than 100 mm, a beam width of less than 1 mm will produce negligible measurement errors. The coil set problem applies also for the third method, the measurement of the first natural frequency.

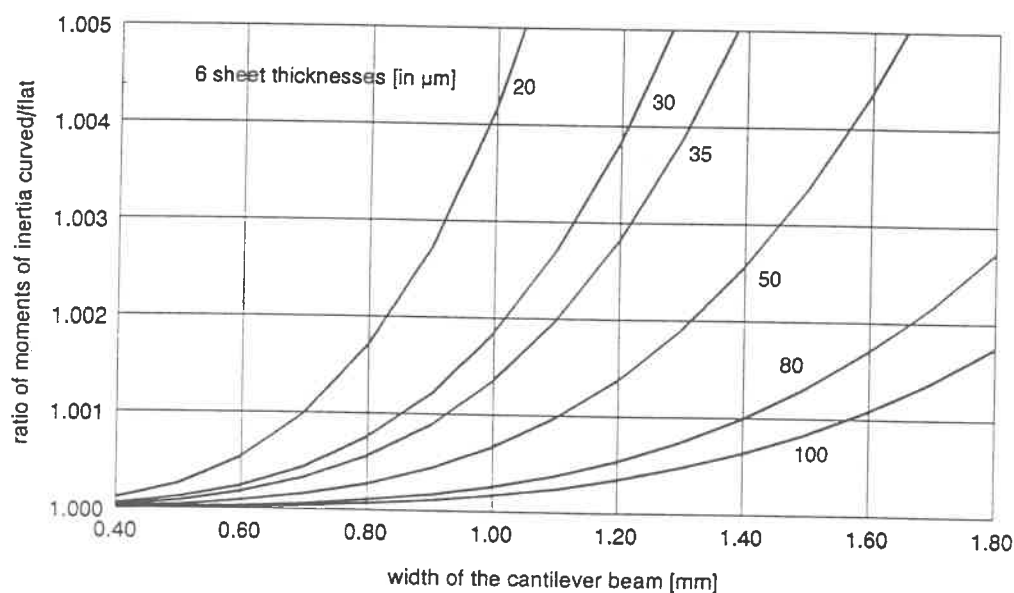


Fig.1. Influence of the coil set of thin rolled sheets at the moment of inertia of cantilever beams. With a coil set of radius $r > 100$ mm and beam width of ≤ 1 mm, the ratio of both moments is ≤ 1.004 .

Measurement of the first natural frequency

The most promising method to determine the non-isotropy of Young's modulus of elasticity proved to be the measurement of the first natural frequency of long, specially designed cantilever springs (Figure 2). Their length axis was placed in the two main sheet directions, in the rolled direction and perpendicular.

The measurement of the first natural frequency of the cantilever springs has been performed for three different spring lengths and for six different foil thicknesses, all in both sheet directions. To obtain reliable results for each type of spring, five identical specimens were used for the measurement. For each specimen six successive values of natural frequency were recorded. A dynamic signal analyser along with electromagnetic initiation of the spring vibration has been used for these measurements.

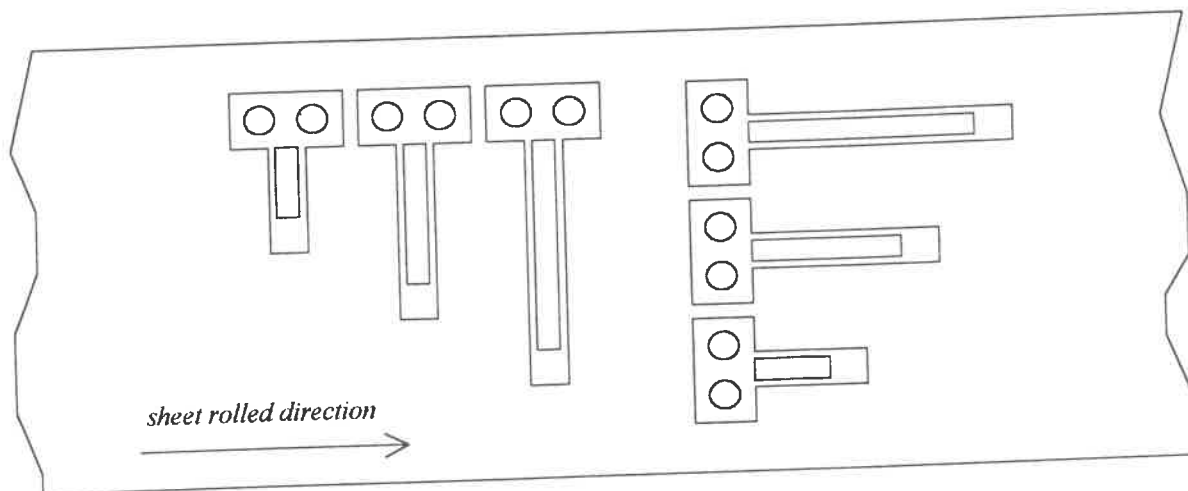


Fig.2. The actual size and orientation of the thin cantilever beam specimens used in the natural frequency measurement method.

Measurement results

The large amount of collected data was reduced by statistical methods to a small number of relevant values. The most reliable is the *ratio* of the two Young's moduli for the two main directions. These six values for each thickness are shown in Figure 3 along with the results of the previous method of bending deflection. With the latter method, three additional thicknesses were taken into consideration including a very thin sheet ($15\text{ }\mu\text{m}$) and a rather thick sheet ($150\text{ }\mu\text{m}$).

Examining the results, the overall trend is very obvious. The material structure following the cold rolled process shows a considerable non-isotropy of Young's modulus of elasticity. The value of the modulus in the direction perpendicular to the rolled direction is generally 1.2 times larger than the modulus value in the rolled direction. The thickness of the sheet plays an important role for the moduli ratio. The thinner the sheet the larger the difference of the Young's moduli is. There is also a remarkable correlation between the moduli ratio and the maximum tensile strength of the particular sheets. This correlation can clearly be seen in figure 3.

To help us understand Young's modulus non-isotropy, an extensive metallurgical analysis of the sheet specimens will be carried in the near future. The first investigations show very long ($50\text{ }\mu\text{m}$) and very thin ($5\text{ }\mu\text{m}$) crystals in the rolled direction. For the thin metal sheets it means that no more than ten crystals fill the entire sheet thickness.

Theoretical calculations

To support the investigations and to verify the measurement results, finite element modal analysis and two different theoretical calculations, effective dynamic mass method and transfer matrix analytical method, were performed to determine the first natural frequency.

These calculations are also necessary to determine the absolute value of Young's modulus of elasticity from the natural frequency measurements. The three calculation methods give slightly different results. The differences are up to 4%. However from our experience with ALGOR® finite element analysis

it is known that this method gives results with an accuracy on the average of 0.5%. Therefore these results are considered most reliable.

By using the finite element results, the average value of Young's modulus in the rolled direction is about 165 GPa and in the perpendicular direction 205 GPa.

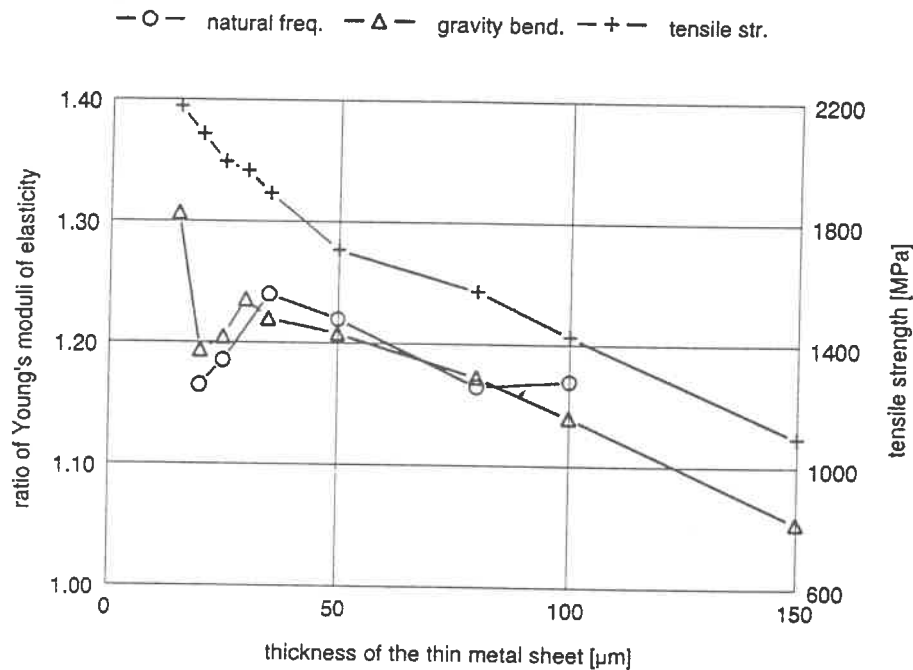


Fig.3. The results of the non isotropy measurements. Young's modulus in the direction perpendicular to the rolled direction is on the average of 1.2 times larger than in the rolled direction. The ratio of both moduli determined by the natural frequency method and the gravity bending method shows a strong similarity. There is a surprising correlation between Young's modulus ratio and the maximum tensile strength of the sheet material.

Conclusions

Accurate measurements of Young's modulus of elasticity of thin stainless steel sheets were performed using two different measurement methods. Differences in the modulus values up to 30% can be seen between the two main sheet directions, in the rolled direction and perpendicular to it. There is also a remarkable correlation between the moduli ratio and the maximum tensile strength of the sheet material.

References

- [1] P.V. Pistecky, H.F. van Beek, H. van der Wulp: Development and Manufacture of Collimators for Gamma Ray Imaging, Precision Engineering Journal of the American Society for Precision Engineering, October 1994, Butterworth-Heinemann, Newton, MA, USA. ISSN 0141-6359.
- [2] P.V. Pistecky, E. van Dop: Design of an Ultra Sensitive Force Transducer for Biomedical Experiments, Proceedings 9th Annual Meeting of the American Society of Precision Engineering, Oct. 2 - 8, 1994, Cincinnati, USA.

Acknowledgements

The author would like to thank Mr. Pavel Puchmajer from the Czech Technical University of Prague for his theoretical calculations and his stimulating discussion of the measurement results.

HIGH PRECISION X-Y- θ MICROPOSITIONING STAGE USING MONOLITHIC FLEXURE-PIVOTED LINKAGES

Jae-Wook Ryu, Dae-Gab Gweon
Dept. of Mechanical Eng., KAIST
Kusong-Dong, Yuseong-Gu, Taejeon, 305-701, Korea

A high precision X-Y- θ micropositioning stage has been developed as an aligner for optical reduction projection and electron beam exposure in lithography system. The micropositioning stage is designed to correct XY axes positioning errors and yaw error due to coarse positioning. In order to achieve large rotation about the θ -axis, the designed stage uses a special monolithic flexure pivoted mechanism. Three piezoelectric actuators with flexure-pivoted mechanical lever system amplifies the rotation range effectively. In addition, to improve immunity to temperature gradients, the micropositioning stage is designed rationally symmetric about the Z-axis. Open-loop characteristics of the micropositioning stage were measured by capacitance displacement transducers. Experimental results indicate that the micropositioning stage has 0.025 μm positioning accuracy over the total range of 38 μm and 30 μm along the X and Y axes, respectively. It also has 0.2 arcsec yaw accuracy over the total range of 350 arcsec about the θ -axis.

Introduction

Precise positioning is an essential technology for step and repeat lithography, particularly for optical reduction projection aligners and electron beam exposure systems. In such systems, a dual stage system consisting of a coarse and a fine table is used in order to achieve fast and accurate positioning.

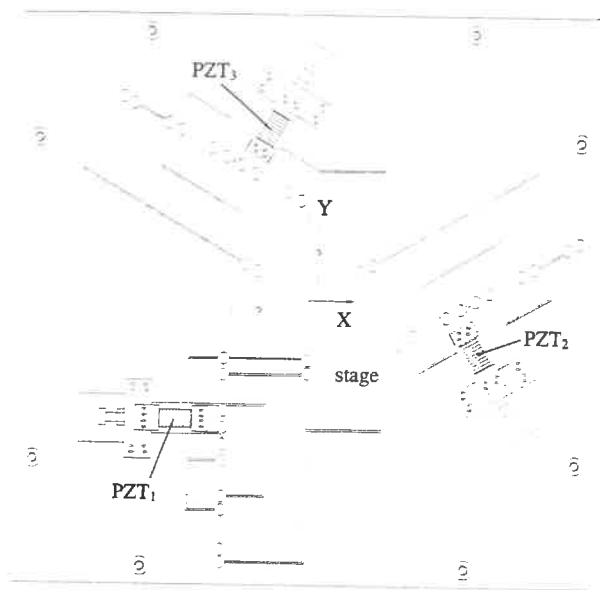
Micropositioning stages utilizing piezo-driven flexure-pivoted linkages can have many advantages: negligible backlash and stickslip friction; smooth and continuous displacement; adequate for magnifying the output displacement of piezoelectric actuators; and inherently infinite resolution. Therefore, they are widely used in many different applications, such as scanning tunneling microscopes (STM)¹⁾, step and repeat lithography²⁾, optical and electron microscopes³⁾, laser welding applications⁴⁾, etc. In this paper, a micropositioning stage is presented primarily designed for a fine table which moves precisely along the X, Y, and θ axes so as to correct positioning and rotation errors from a coarse table positioning. Particularly the yaw error from the

coarse table is amplified as traveling along the XY axes, the proposed stage has large rotation capacity about the θ -axis.

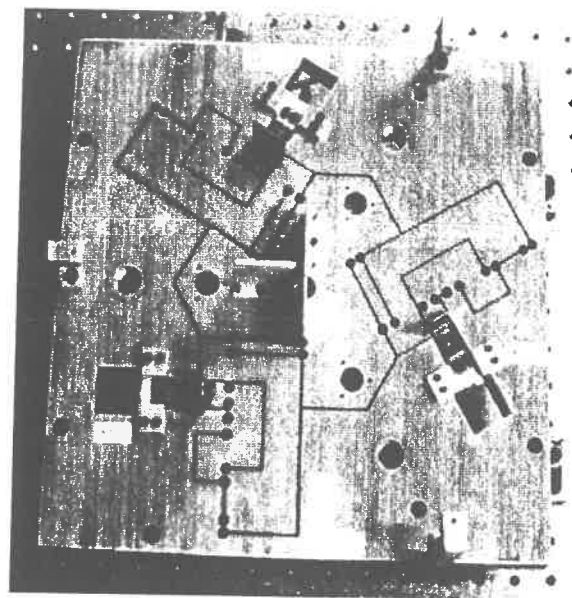
In this paper, it is shown that the precision X-Y- θ micropositioning stage has a much larger rotation angle than parallel linkage type micropositioning stages²⁾. In addition, some experimental results of the newly developed micropositioning stage are presented.

Micropositioning stage design

The stage design combines three piezoelectric actuators and monolithic flexure-pivoted linkages. To achieve a large rotation angle about the θ -axis, three input mechanism amplifies length change of a piezoelectric actuator by using flexure-pivoted mechanical lever system. Each one of the input mechanisms use a double compound flexure-pivoted lever system in order to minimize the lever length and maximize the mechanical displacement gain (Fig. 1). In addition, to improve immunity to temperature gradients, the stage is designed rationally symmetric about the Z-axis by placing each input mechanism 120 degrees apart



(a)



(b)

Fig.1 X-Y- θ micropositioning stage

about the Z-axis. The monolithic flexure-pivoted linkages are composed of typical notch flexure hinges which was machined from a $320 \times 320 \times 20$ mm metal blank. All flexure hinges have a hinge radius of 3.3 mm and a width of 20 mm. However, to maximize the rotation angle, minimum thicknesses of the hinge are varied from 0.5 mm to 2.9 mm while taking account of the allowable stress and the mechanical displacement gain of the input mechanism⁵⁾. The stage was built from 7075-T6 aluminum alloy.

Each piezoelectric actuator is placed in a cutout actuator housing and preloaded with an axial force by pushing a bolt in order to be free from backlash. The piezoelectric actuator has dimensions of 10 by 10 by 18 mm and the rated maximum output displacement of $15 \mu\text{m}/100\text{V}$.

Experimental system configuration

Figure 2 shows the block diagram of the experimental system. The each command voltage from PC was converted to analog (0~10V) by a 12 bit D/A. And then the voltage was amplified approximately 10 times by using a power amplifier. Each piezoelectric actuator was driven independently. The open-loop displacement characteristics of the stage were measured with the use of six capacitance gages (Model 3890, ADE

corp.) with $0.025 \mu\text{m}$ resolution. The relationship between input and output displacements for the stage was obtained by measuring the output displacements of the stage, translations along the X and Y-axis and rotation about the θ -axis, and the length changes of the piezoelectric actuators. Three gages, s_1 , s_2 , and s_3 , were used in order to measure the output displacements of the stage. The other gages, s_4 , s_5 , and s_6 , were used in order to measure the length changes of the piezoelectric actuators (Fig. 3).

Kinematic calibration

The relationship between input and output displacements for the stage should be known in order to move the stage to the desired position. From the some sets of measured input and output displacements, the relationship is obtained by using the following manner.

Suppose the relation is written in the form:

$$\mathbf{x}_i = \mathbf{T}_{io} \mathbf{x}_o \quad (1)$$

where, $\mathbf{x}_i = [u_1, u_2, u_3]^T$ are the length changes of the piezoelectric actuators, $\mathbf{x}_o = [X, Y, \theta]^T$ are the output displacements of the stage, and $\mathbf{T}_{io}$ is the transformation matrix (3×3) from the output displacements to the input

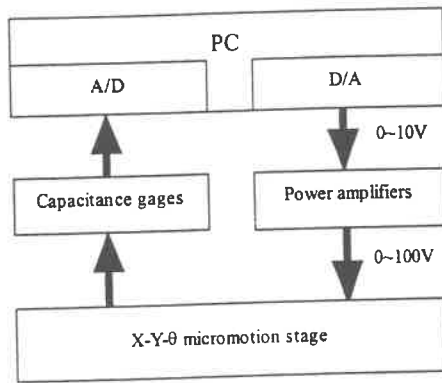


Fig.2 Block diagram of the stage system

displacements, respectively. From Eq. (1), the n sets of measured input and output displacements should satisfy the following relation:

$$\begin{matrix} X_i \\ (3 \times n) \end{matrix} = \begin{matrix} T_{io} \\ (3 \times 3) \end{matrix} \begin{matrix} X_o \\ (3 \times n) \end{matrix} \quad (2)$$

$$\text{where, } \begin{matrix} X_i = [x_i^1, x_i^2, \dots, x_i^n] \\ X_o = [x_o^1, x_o^2, \dots, x_o^n] \end{matrix}$$

From Eq. (2), T_{io} can be found from the measured results of X_i and X_o experimentally.

$$T_{io} = X_i X_o^T (X_o X_o^T)^{-1} \quad (3)$$

From the obtained T_{io} , we can find the required length changes of the piezoelectric actuators to move the stage to the desired position.

In like manner, the transformation matrix from the input displacements to the command input voltages are obtained. That is,

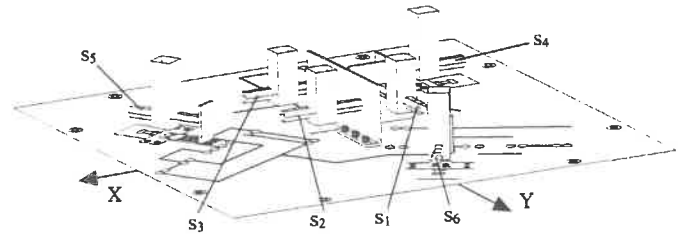


Fig.3 Schematic diagram of measurement system

$$T_{vi} = V X_i^T (X_i X_i^T)^{-1} \quad (4)$$

$$V = [v^1, v^2, \dots, v^n]$$

where, $v = [v_1, v_2, v_3]^T$ are the command input voltages from PC to D/A.

Experimental results and discussions

Figure 4 shows the length changes of the piezoelectric actuators and the corresponding output displacements of the stage, when the command input voltages for driving piezoelectric actuators, PZT₁, PZT₂, and PZT₃, are all the same. At this driving condition, ideally, it is expected that the length changes of the piezoelectric actuators, u_1 , u_2 , and u_3 , are all the same. Therefore, the stage will rotate about the θ -axis only. However, as can be seen from Fig. 4, experimental results indicate that the length changes of the piezoelectric actuators are somewhat different. Therefore, position errors, X and Y translations, are introduced. The different

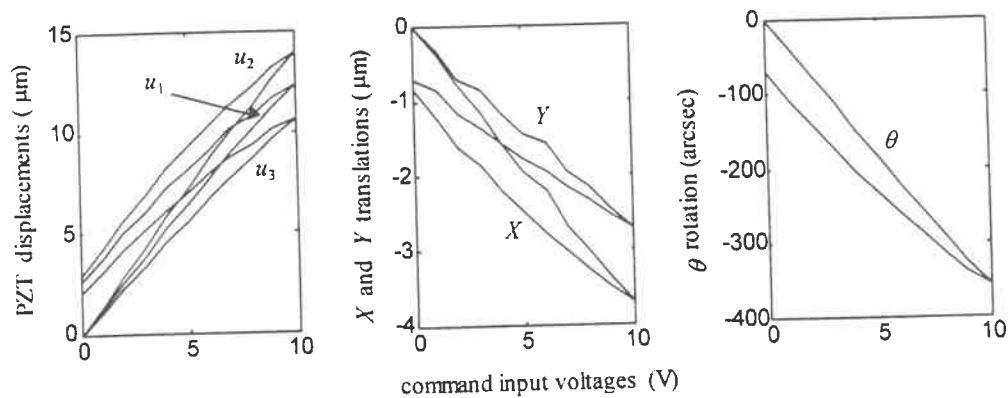


Fig.4 Length changes of the piezoelectric actuators and the output displacements of the stage, when the command input voltages for driving PZTs are all the same

length changes of the piezoelectric actuators are due to the manufacturing tolerance of the stage, the different preload and rated maximum output displacement of each piezoelectric actuator, and the different electrical gain of each power amplifier. The hysteresis of piezoelectric actuator is shown in Fig. 4.

The interference characteristics of each input are shown in Fig. 5. When PZT₁ is driven alone, small values of u_2 and u_3 are occurred in proportion to u_1 . Because the stage is symmetric, similar results are obtained when PZT₂ and PZT₃ is driven alone, respectively. The interference ratio is below 5%. This means that each piezoelectric actuator is free from the other two piezoelectric actuators.

The open-loop step responses of the stage are shown in Fig. 6 and 7. If the desired output position of the stage is given, the command input voltages are calculated using T_{io} and T_{vi} , which had been found from Eqs. (3) and (4), respectively. Then, the calculated input voltages are commanded from PC, the stage will move to

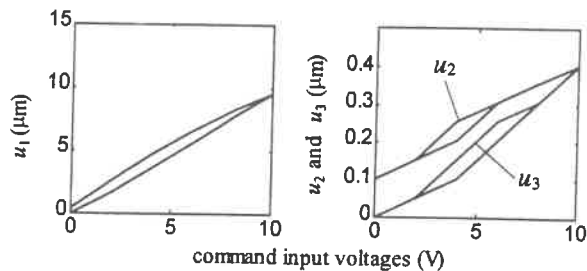


Fig.5 Input interference

the desired position with some open-loop errors. In order to verify the input-output relationship of the stage, step responses of the stage are experimented along the X, Y, and θ -axis, respectively. As can be seen from Fig. 6, the stage can be moved along the desired direction with small parasitic errors. Figure 7 shows open-loop step responses, when the desired step sizes are the resolution of $0.025 \mu\text{m}$ along the X and Y-axis and 0.2 arcsec about the θ -axis, respectively.

Table 1 shows the total ranges of output

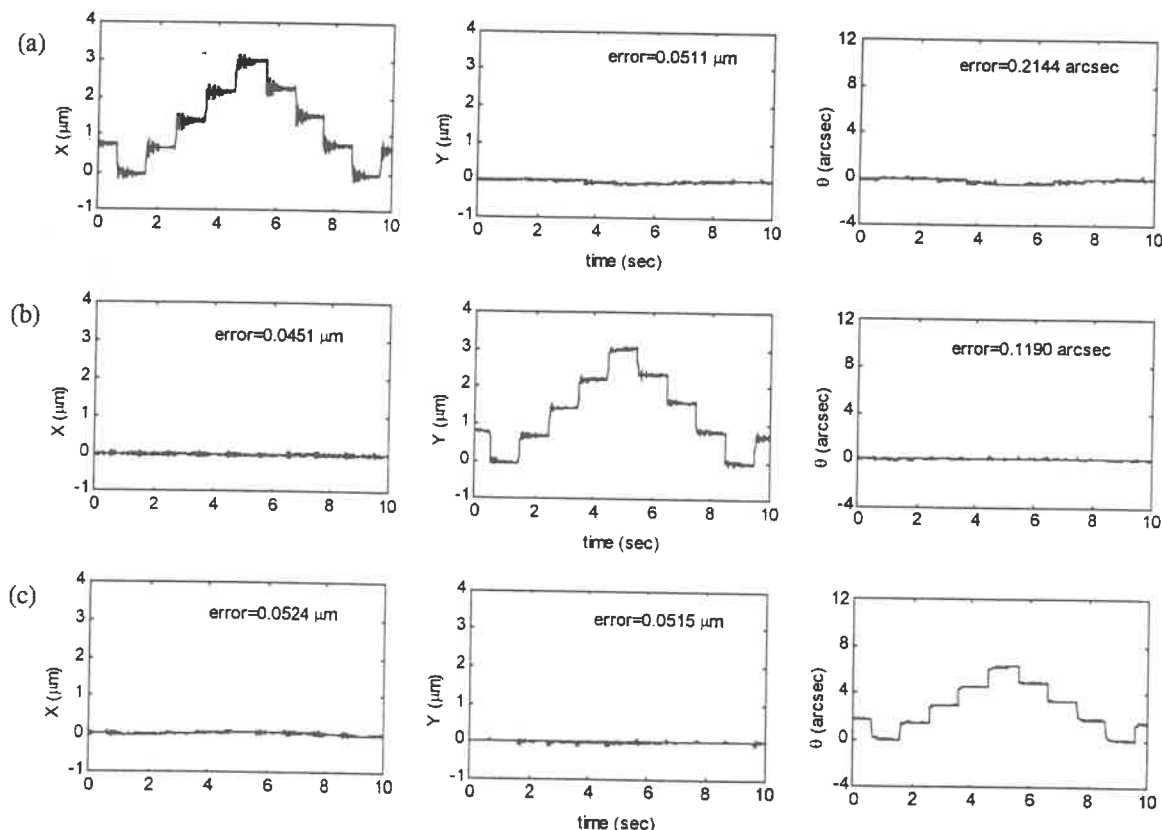


Fig.6 Step responses of the stage along (a) X, (b) Y, and (c) θ -axis (open-loop control)

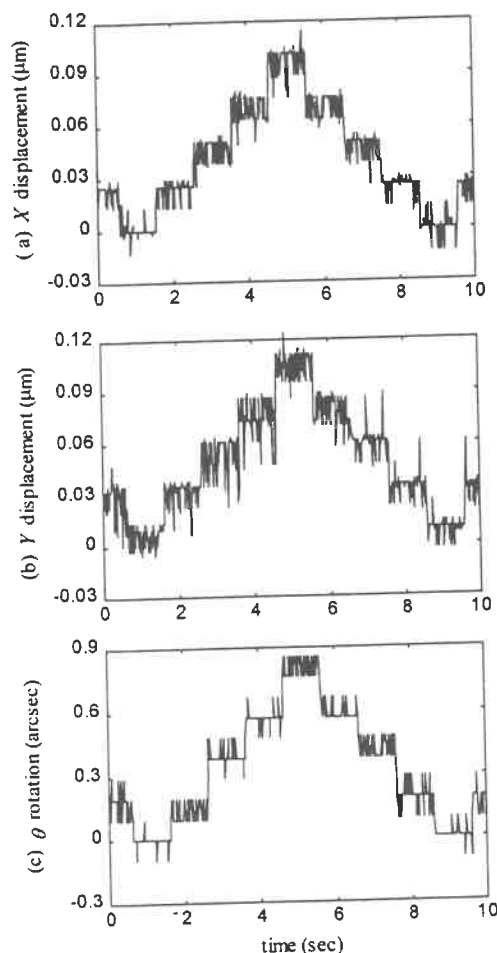


Fig.7 Resolution of the stage

displacements and the resolution along the X, Y, and θ -axis. The stage has $0.025 \mu\text{m}$ positioning accuracy over the total range of $38 \mu\text{m}$ and $30 \mu\text{m}$ along the X and Y-axis, respectively, and 0.2 arcsec yaw accuracy over the total range of 350 arcsec about the θ -axis. Especially, the total range about the θ -axis is more than 10 times larger than parallel linkage type micropositioning stages²⁾.

Conclusions

In this study, a high precision X-Y- θ micropositioning stage has been developed for use of a fine table which corrects the XY axes positioning errors and the yaw error due to a coarse table and its open-loop characteristics have been experimented.

The developed micropositioning stage has $0.025 \mu\text{m}$ positioning accuracy over the total range

Table 1 The total ranges of output displacements and the resolutions

		ranges	resolutions
X	(μm)	38	0.025
Y	(μm)	30	0.025
θ	(arcsec)	350	0.2

of $38 \mu\text{m}$ and $30 \mu\text{m}$ along the X and Y-axis, respectively. It also has 0.2 arcsec yaw accuracy over the total range of 350 arcsec about the θ -axis, which is about 10 times larger than that of the parallel type micropositioning stages.

Acknowledgments

This work was funded by Korea Science Foundation (KOSEF) under grant 94-0200-09-01-3. The authors wish to thank KOSEF for their financial support.

References

- 1) E. Furukawa, M. Mizuno, and T. Hojo, "A Twin-Type Piezo-Driven Translation Mechanism", *Int. J. Japan Soc. Prec. Eng.*, Vol. 28, No. 1, Mar. 1994, pp. 70-75.
- 2) S. Moriyama, T. Harada, and A. Takanashi, "Precision X-Y Stage with a Piezo-driven Fine-table", *Bull. Japan Soc. of Prec. Engg.*, Vol. 22, No. 1, Mar. 1988, pp. 13-17.
- 3) F. E. Scire and E. C. Teague, "Piezodriven $50\text{-}\mu\text{m}$ range stage with subnanometer resolution", *Rev. Sci. Instrum.*, Vol.49, No.12, Dec. 1978, pp. 1735-1740.
- 4) R. Yang, M. Jouaneh, and R. Schweizer, "Design and characterization of a low-profile micropositioning stage", *Prec. Eng.*, Vol. 18, No. 1, Jan. 1996, pp. 20-29.
- 5) D. G. Gweon, *Development of Ultraprecision Positioning System Using PZT Actuators*, KOSEF research report, KAIST, 1996 (in Korean).

Innovations in the design and development of ultraprecision machines

Lightweight and ultrastiff single point diamond turning machine with submicron machining accuracy

J.P.M.B. Vermeulen, P.C.J.N. Rosielle, P.H.J. Schellekens

Precision Engineering Group

Department of Mechanical Engineering

Eindhoven University of Technology (the Netherlands)

Abstract

Main structural components of most existing single point diamond turning (SPDT) machines are made of granite, and conventionally, steel is applied in hydrostatic slideways. The structural design of moving parts of SPDT machines may be substantially improved, with regard to stiffness and dynamic behavior, using ceramic laminate materials with high specific stiffness (E/ρ). Applying readily-obtainable alumina ceramic tiles in a plate frame type SPDT machine - which is currently under construction - the slide's mass is reduced with a factor of three compared to an equal slide construction of steel or aluminum with the same stiffness.

A statically determinate and thermomechanically stable machine design significantly improves reproducibility. As a result, additional software compensation of residual errors, necessary to meet currently required submicron machining accuracy, is limited.

Based on existing compensation techniques, an infinitely stiff, reproducible aerostatic thrust bearing was developed. Direct drive systems are used for the slides and the main spindle, in order to eliminate a series of mechanical compliances as well as friction forces and hysteresis.

Innovations in this design are not confined to SPDT machines alone, but can readily be transferred to ultraprecision machines and instruments more generally.

Keywords: Ultraprecision machines, ceramic laminate materials, infinitely stiff aerostatic thrust bearings, direct drive systems

Introduction

In line with advances in precision manufacture, accuracies required in new industrial products continue to be demanding. Ultraprecision parts in particular, such as optical mirrors, molds for injection molding, and aerostatic bearings, are realized outstandingly through single point diamond turning (SPDT). In addition to non-ferrous metals and plastics, recently ferrous metals were machined with single point diamond tools, using an ultrasonic cutting device.

As a result of the need for extension of the ultra precision machining capacity in the Eindhoven University's Engineering and Design Facilities, an SPDT machine was designed by the

Eindhoven Precision Engineering Group [Vermeulen, 1994]. The machine, that is currently under construction, differs significantly from conventional SPDT machines.

Demands on slide constructions in high precision machines can be derived from the stated positioning accuracy. Single point diamond turning however, is more than a point-to-point positioning process. Each deviation from the diamond tool's desired movement, for example as a result of finite stiffness or thermal instability of the machine, will be directly depicted on the workpiece.

This paper briefly discusses three topics in our SPDT machine design, which highly determine its overall behavior. After a short description of the machine, the application of ceramic laminate materials, aerostatic thrust bearings and direct drive systems is highlighted.

SPDT machine design

Machining accuracy is closely related to predictability and especially reproducibility. The most cost-effective way to achieve this, is to create a statically determinate and thermomechanically stable machine design. In this way, minimum tolerances in the manufacture of the machine as well as minimum demands on conditioning of ambient temperature are required. The objective was to attain accuracies in form and size of better than $0.1 \mu\text{m}$ over 100 mm, by a sound mechanical design. The effort to set up software compensation is significantly reduced, for a system with good inherent accuracy.

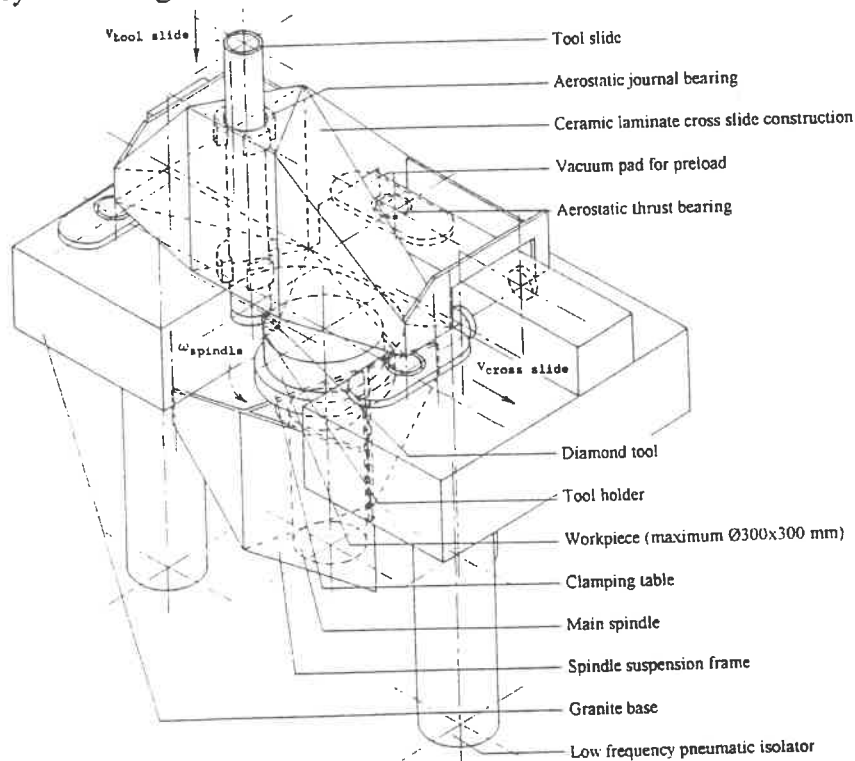


Figure 1. Schematic representation of lightweight and ultrastiff SPDT machine, currently under construction.

In figure 1, a schematic representation of the SPDT machine is shown. A turret configuration was chosen, since workpieces up to $\varnothing 300 \times 300$ mm cause an inadmissible deflection when mounted in a classical, horizontal set-up. The cross slide, that consists of a plate frame structure, runs on a granite guide way. Using a plate frame type slide construction, the utilization ratio of the applied material is optimized, giving a very favorable force transmission, and therefore, a lightweight and stiff design. Underneath the guide way, the main spindle is mounted in a triangular suspension frame. Thereby, the centerline of the workpiece remains stable, independent of room temperature fluctuations. The granite base is supported by three low-frequency pneumatic isolators (1 Hz) in order to eliminate disturbing vibrations. Automatic height-leveling systems are built in for varying load conditions (slide travel).

Ceramic laminate slide construction

When machining non-ferrous metals and plastics, diamond turning forces are very small [Oomen, 1992]. Therefore, in the design of an SPDT machine's slide construction, it is the material's modulus of elasticity (stiffness) rather than the admissible stress, on account of which a material should be selected.

A considerable mass of a slide construction causes high inertial forces as well as a substantial heat dissipation in the slide's driving system. Another reason to reduce the slide's mass is the increase of its natural frequency $\omega_e = \sqrt{c/m}$.

For plate frame type slide structures, both mass and stiffness are linear proportional to the sheet thickness. Therefore, optimizing the stiffness per kg (c/m) for a certain geometry constitutes to select a material with a high specific stiffness E/ρ [m^2/s^2]. In slide constructions of most existing ultra precision machines, conventional construction materials, like steel, aluminum and granite are applied. For the high specific stiffness of massive ceramics (factor 3.1 to 6.7 times E/ρ of steel, aluminum or granite), applying these might reduce the slides's mass significantly. The obtainability however, of standardized ceramic parts is limited. Only for Al_2O_3 -ceramic, plates and sheets are standard obtainable and therefore priced reasonably.

Vibrations, caused by slide movements or coming from the environment, will influence the slide's position accuracy. Whilst the specific stiffness of massive ceramics is good, its vibration damping is bad. Therefore, in ceramic structures for precision engineering applications, a laminated build-up is highly preferable. The glued joints are barriers for vibrations, so that vibration damping is increased considerably.

Another advantage of a laminated build-up is the greater designer's freedom opposite to massive ceramic structures. Depending on the way of manufacturing, ceramic parts may have many internal defects. They also show a lot of shrinkage and deformation, in consequence of transformations at high temperature sintering. Therefore, in the design of massive ceramic parts, a uniform sheet thickness, relatively simple structures and the absence of sharp edges, and point or line forces are prescribed. For alumina ceramic plates with dimensions of typically 500×500 mm a minimum thickness of 25-50 mm is necessary. Complex ceramic plate frame structures with sharp included angles can therefore better be build out of prefabricated laminated plates. The Al_2O_3 -ceramic laminate material has been patent applied for [Rosielle, 1996].

Experiments on a 5-layer Al_2O_3 -ceramic laminate tensile strip, displayed an E/ρ factor of $75 \cdot 10^6 \text{ m}^2/\text{s}^2$. That is a factor of 2.9 and 2.8 more than an aluminum or steel tensile strip respectively with equal mass.

In parts made of a homogeneous and isotropic material, the utilization ratio of the applied material μ [-] characterizes the stress distribution over the particular part [Hoek, vd, 1989]. If the stress is equal within the entire material, for example in a uniformly loaded massive tensile strip, $\mu=1$, and for bending of a sheet of rectangular section by a constant moment, $\mu=1/3$.

From calculations on the ceramic laminate material, the number μ appeared to depend on the number of layers N . The in-plane stiffness of a multilayer ceramic laminate structure (N large), approaches the stiffness of the massive ceramic material [Vermeulen, 1996].

At present, the SPDT machine's cross slide construction is being build-up out of alumina ceramic laminates.

Aerostatic thrust bearings

Next to hydrostatic bearings, aerostatic bearings are commonly used in ultraprecision machines for their negligible friction and hysteresis. Until now, as air is highly compressible in comparison with oil, aerostatic bearings determine the overall obtainable machine stiffness.

In literature, a lot of investigation work on aerostatic bearings can be found. However, as different values of bearing parameters were chosen in different publications, it is very difficult to reserve a fair judgment about specific bearing features among different types.

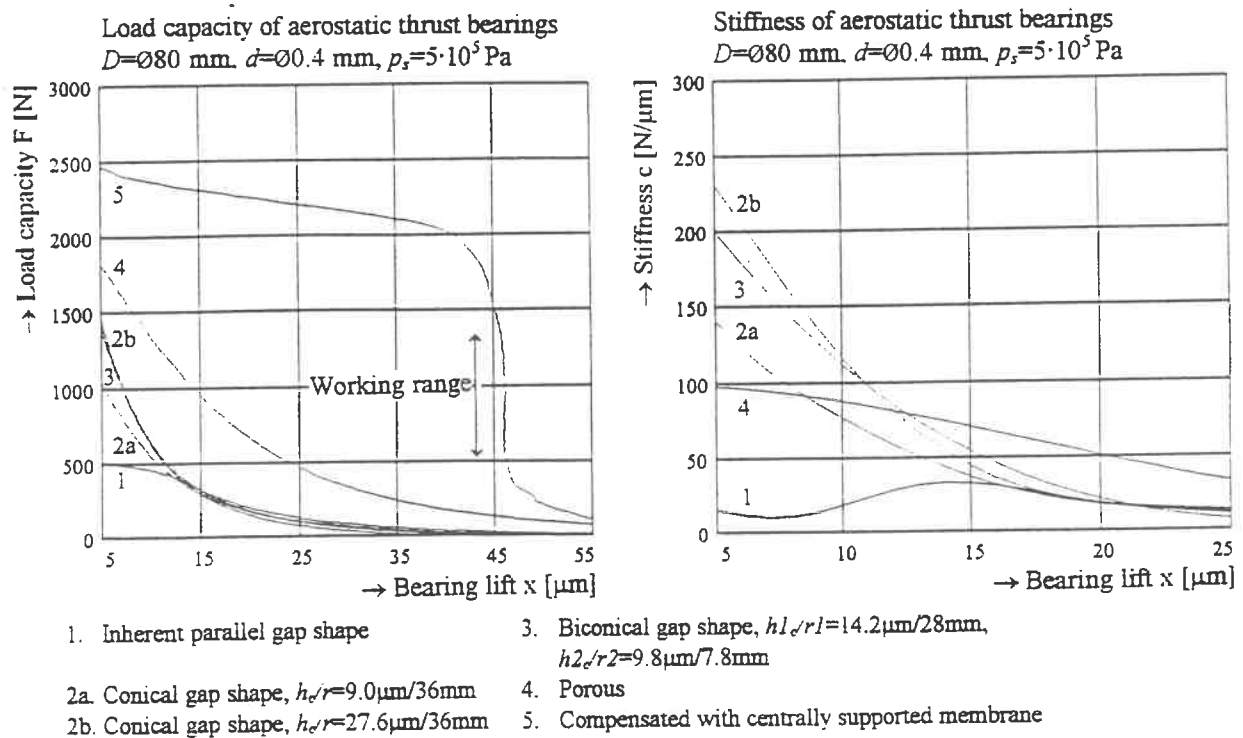


Figure 2a. Measured load capacity of aerostatic thrust bearings.

Figure 2b. Calculated stiffness of aerostatic thrust bearings.

Therefore, five different aerostatic thrust bearings were made in-house, tested and mutually compared. The main proposal was to get a survey of stiffness and load capacity. Next to conventional parallel and conical gap shape bearings, biconical as well as porous bearings were investigated. The bearings all have an outer diameter $D=80$ mm, a central restrictor with $d=0.4$ mm (except for the porous bearing), and are supplied to an equal pressure $p_s=5 \cdot 10^5$ Pa. A compensated aerostatic bearing with centrally supported membrane was made as well. Here, the bearing gap geometry is used as a controlling element. The convergence of the gap in the flow direction is forced to become an increasing function of the external load. Provided that the membrane deflection equals exactly the bearing displacement as a result of the compression of air in the bearing gap, an infinite bearing stiffness can be obtained [Blondeel, 1976], [Bryant, 1986], [Snoeys, 1987].

From the measured load capacity curves (about 500 data points), shown in figure 2a, the stiffness curves are calculated by means of a polynomial curve fit (see figure 2b). Results are conveyed in table 1. As the flatness of a guiding surface is typically of the order of $1 \mu\text{m}$, a minimum air gap of $5 \mu\text{m}$ is used.

Bearing type $D=\varnothing 80$ mm $d=\varnothing 0.4$ mm $p_s=5 \cdot 10^5$ Pa	Remarks	Maximum stiffness [N/ μm]	Air gap for maximum stiffness [μm]	Load capacity for maximum stiffness [kN]	Maximum load capacity [kN]
Inherent parallel gap shape		32	14	0.3	0.5
Conical gap shape	$h_c/r=9.0\mu\text{m}/36\text{mm}$	136	5	1.0	1.0
	$h_c/r=27.6\mu\text{m}/36\text{mm}$	230	5	1.5	1.5
Biconical gap shape	$h1_c/r1=14.2\mu\text{m}/28\text{mm}$	200	5	1.4	1.4
	$h2_c/r2=9.8\mu\text{m}/7.8\text{mm}$				
Porous	Graphite UT6ST porosity 17%	99	5	1.8	1.8
Compensated with centrally supported membrane		∞	5-15	0.5-1.3	1.5

Table 1. Stiffness and load capacity of investigated aerostatic bearings.

For the infinite stiffness of the compensated bearing, a reproducible design was made and utilised in the SPDT machine.

Direct drive systems

In high precision machines, a wide range of drive systems is applied. Lead screw drives have considerable friction and hysteresis. As a result of the over determination, the accuracy of a lead screw drive is determined by a statistic mean of many contacts. Friction wheel drives in contrary, can have a prescribed level of preload at the wheel contacts, and therefore, its behavior is more predictable and reproducible.

Linear direct drive systems have no contacts and therefore no friction at all. A series of mechanical compliances is eliminated. The drive's stiffness is determined in the closed loop

servo system and depends directly on the resolution of the measurement system. As the application of linear direct drive systems in combination with linear optical encoders in SPDT machines is relatively new, its smoothness of motion was tested at different speed rates. Using sinusoidal software commutation and position feedback with 5 nm resolution zerodur scales, the slide's position error signal is restricted to about 0.1 μm . Heat dissipation in the applied linear motor is less than 1 W.

Next to the cross slide and the tool slide, a direct rotational drive system is applied for the main spindle. For its very small ripple torque as a result of the great number of turns, a disk armature coil motor is used. In order to eliminate the noise, originating from the motor's ball bearings, the latter were removed. Instead, the rotor is fixed directly to the spindle, which has an aerostatic bearing system.

Conclusions

Machining accuracy highly depends on stiffness, reproducibility and stability, keypoints in the development of a new, universal SPDT machine.

Applying alumina ceramic laminate materials, the slide's stiffness to mass ratio is increased with a factor of three, compared to conventional construction materials. Infinitely stiff aerostatic thrust bearings eliminate the determining compliance of the machine, and significantly reduce the slides' friction and hysteresis. The latter two are further reduced, using direct drive systems.

As a result, the amount of additional software compensation, necessary to meet submicron machining accuracy, is significantly reduced.

References

- Blondeel, E., Snoeys, R., Devrieze, L., 1976, Aerostatic bearings with infinite stiffness, In: *Annals of the CIRP*, 25(1).
- Bryant, M.R., Velinsky, S.A., Beachley, N.H., Fronczak, F.J., 1986, A design methodology for obtaining infinite stiffness in an aerostatic thrust bearing, In: *Journal of mechanisms, transmissions, and automation in design*, Transactions of the ASME, 86-DET-32.
- Hoek, W. v.d., 1989, A mechanical designers case book, Philips CFT Report 59/85, translation of 'Des Duivels Pentenboek' (D.D.P.).
- Oomen, J.M., Eisses, J., 1992, Wear of monocrystalline diamond tools during ultraprecision machining of nonferrous metals; In: *Precision Engineering*, 14(4), pp. 206-218.
- Rosielle, P.C.J.N., Vermeulen, J.P.M.B., Reker, E.A.G., 1996, Laminaten, vervaardigd uit epoxy-verlijmde, dunne Al_2O_3 -keramische tegels, geschikt voor sledeconstructies in precisietoepassingen (Laminates, made of epoxy glued, thin Al_2O_3 -ceramic tiles, suitable for slide constructions in high precision applications), License request no. 1002369 (in Dutch).
- Snoeys, R., Al-Bender, F., 1987, Development of improved externally pressurized gas bearings, *KSME Journal*, 1(1), pp. 81-88.
- Vermeulen, J.P.M.B., 1994, Een ontwerp van een diamantdraaibank met submicron bewerkings-onnauwkeurigheid, (Design of a diamond turning machine with submicron machining inaccuracy), Masters thesis (in Dutch), Eindhoven University of Technology (the Netherlands).
- Vermeulen, J.P.M.B., 1996, Development of a lightweight and ultrastiff single point diamond turning machine with submicron machining accuracy, Mechatronics post-masters design thesis, Eindhoven University of Technology (the Netherlands).

Straightness Errors of Rectangular Beams Caused by Ambient Air Temperature Gradients

Kevin L. Wasson¹
Alexander H. Slocum²
John H. Lienhard V³

In a typical industrial building, vertical temperature gradients within the air may be 2°C per meter or greater, due to the tendency of air to stratify. Consequently, a beam with its length oriented horizontally will experience warping as its top becomes warmer than its bottom. The degree of warpage will depend upon the beam geometry, how it is fixtured, and its material properties. Our calculations show that this effect is important in applications such as coordinate measuring machines where micron-level precision is required, particularly when ceramic materials are used because they have very low thermal conductivities. Our work investigates the specific case of thermal straightness errors of rectangular beams caused by ambient air temperature gradients. Numerical simulations of heat conduction are used to evaluate the effect of beam geometry and material properties on beam warpage. These simulations are performed using experimentally-determined heat transfer coefficients applied to the beam surfaces and calculating the two-dimensional temperature field. Transient simulations are used to investigate beam sensitivity to changes in ambient temperature. Case studies are presented which illustrate the effect of beam material on thermal stability. Our results show that beams made of low-conductivity, low-expansivity ceramics can have considerably higher steady-state thermal deflections than similar beams made of high-conductivity, high-expansivity metals. Ceramic beams are less sensitive to fluctuations in ambient temperatures, however, because their thermal time constants are generally very high. The simulations are also used to investigate the potential advantages of using various insulating layers of materials to thermally isolate beams from their surroundings, and reduce thermal errors. A novel method is presented in which a layer of insulation followed by a thin layer of high-conductivity metal such as aluminum is used to reduce the thermal warpage by over an order of magnitude. Although this work presents only numerical simulations that have not been verified with experimentation, we believe that the work is valuable at least qualitatively because it improves our understanding of the thermal problems that are present in precision machines. In particular, the results representatively demonstrate that using materials with lower thermal expansivities does not always lead to lower thermal errors.

Key Words: thermal errors, thermal stability, ceramics

¹Aesop, Inc. Bow, NH 03304.

²Alex & Britt d'Arbeloff Professor of Mechanical Engineering, M.I.T. Precision Engineering Research Group.

³Professor of Mechanical Engineering, Warren M. Rohsenow Heat and Mass Transfer Laboratory, M.I.T.

Using Chebyshev Norms for Estimating Machine Tool Error Parameters

Hooman Tajbakhsh, DOE Graduate Research Fellow
Zainul Abadin, Graduate Research Assistant
Placid M. Ferreira, Associate Professor *

University of Illinois, Department of Mechanical
and Industrial Engineering, Urbana IL 61801

August 28, 1996

Abstract

Numerically controlled machine tools use their structure as a reference frame for monitoring and controlling the position of the cutting tool. The tool positioning accuracy and, eventually, the accuracy of parts machined (if other sources of errors are not overwhelming) are thus dependent on the accuracy of this reference. Manufacturing and assembly are known to leave behind substantial errors in this reference frame and quasistatic effects (thermal and flexural) can cause significant changes in these errors [10, 7, 3, 5, 12, 9, 8]. The problem of characterizing and compensating these errors is thus one of immediate importance.

A good deal of research has been reported on the modeling of these errors [4, 1, 11, 6, 8] and, given an adequately parameterized model, the production of compensated tool motion. Scant attention has been paid to the identification of model parameters and it would not be far from the truth to state that least squares has been the universally used criterion for identifying model parameters from observations. That is, parameters are identified for an error model so that the sum of squares of the difference between the actual machine errors and model predictions are minimized. Thus, for a machine using a model developed with such parameters and visiting a number of points in its workspace, nothing can be said of its accuracy at any one point. All that can be said is that if the sample of points visited represents an unbiased sample of the workspace, the mean squared errors of the machine at these points is minimized.

* Author for Correspondence

A part produced by an NC machine-tool will be evaluated, not on the basis of the average of all points on a machined feature, but on the basis of every point on a machined feature meeting a said tolerance criterion (refer to ANSI Y14.5). Similarly, the accuracy of an inspection task performed on a CMM does not depend on its average accuracy across the workspace but on its accuracy at the point at which it performs the task. Unfortunately, estimating model parameters by using the least squares criterion does nothing to minimize or bound the error or discrepancy between the model and observations at a single point. In fact, it might trade the discrepancy at a small subset of points to obtain a better average performance. The volumetric errors resulting from such compensation are thus unevenly distributed across the workspace and this effect is magnified in the event of the model being inadequate, resulting in a 'sweet spot' in the work volume where the model performs well and large errors outside this small region in the workspace.

One approach to overcome such problems is to choose a different criterion for identifying the model parameters. Estimating model parameters by minimizing the maximum difference between observed errors and those predicted by the model, the classical Chebychev or L_∞ norm [2], takes care of the problems mentioned above. Instead of trading off prediction errors at individual observation points to improve the mean square error, it reins in the large prediction errors, thus attempting to equalize the prediction error across the workspace. An end product of such approach is a tight bound on the maximum prediction error at all observation points and, assuming that the observation set properly characterizes the workspace, a close estimate of the worst case performance across the workspace. From a practical point of view, such a parameter estimation scheme would be important for a) compensation schemes on machines executing different and unpredictable work cycles (and hence using different parts of the workspace in no particular pattern) as those functioning in a FMS and b) acceptance testing of a machine tool, i.e., characterizing its best possible worst volumetric error.

Least squares criterion penalizes the larger errors more than the smaller ones because the error terms are squared. It is mathematically convenient, but it can still leave behind large prediction errors at some observation points if it finds a sufficient number of other observations points to offset this in a mean squared sense. The L_∞ norm also known as the Chebychev's norm is only concerned with the absolute value of the maximum error. While not as convenient in usage as least squares, it is particularly important in establishing worst-case behavior and could prove useful in (i) acceptance testing of machine tools, and (ii) software error compensation of general-purpose machine tools. Given an adequate sample of observation points, in machine tool applications, this criterion could guarantee that the maximum error across the workspace does not exceed some value.

The error model we use in this paper is one developed in this program of research and reported in [8, 10, 7]. The error model describes how the components of the error vector are related to structural errors of the links and transmission errors of the joints of the kinematic scheme of the machine. Using the link and joint transformation model for a serial mechanism, the kinematics of the machine-tool are modeled using a series of alternating fixed and moving co-ordinate frames. Next, each link transformation (which models the constant co-ordinate

transformation between the two joints connected by the link) is given six degrees of freedom to model the errors in the geometry of the represented link. The model contains 17 linearly independent parameters.

The performance of the parameter estimation techniques (least squares and Chebychev's linear program) are compared based on experimental results. Lacking an absolute means of measuring (a sufficiently accurate calibration artifact) errors in a machine tool's workspace, we use the measurements of an artifact in the 'cold state' of the machine as a reference. By taking the machine through a thermal cycle (exercising all its axes), we get the quasistatic effects to change the errors of the machine and observe the ability of our models to track these changes in (instead of the absolute) errors. An artifact with 27 optimally located points along with a 3 DOF probe is used to measure the errors in the x , y , and z directions. The probe has a resolution of 0.25 microns with 1 micron repeatability. A total of eight measurement sets were collected. Each of these measurement sets give rise to 78 independent observed errors at a particular thermal state of the machine. We then use these observations to study the model parameters using least squares and linear programming approach.

The Chebychev's fit shows that the model explains the errors in the workspace under the thermal conditions after 9.56 hrs of heating to within a maximum error of 0.000146 inches while the least squares fit has a maximum value of 0.000186 inches (The maximum observed errors were around 0.0027 inches). Both models perform well but the results reflect about a 25% improvement in terms of the maximum prediction error if the Chebychev's fit is used to estimate the parameters of the error model. It is seen, as expected, that the RMS error for the least squares estimates is better than the Chebychev, but is extremely small or even negligible for both.

In the next linear program, the entire eight set of experiments are used for the constraints of the linear program. The eight sets of experiments give rise to 1248 constraints. These results show that there exists a set of parameters that optimizes for the entire work-cycle (or temperatures) to produce a maximum of 0.000776 inches compared to a 0.000952 inches for the least squares solution. It is again clear that the LP fit provides a better solution in terms of the maximum error than least squares. However, this performance (in terms of lower maximum error) might be inadequate for a compensation scheme.

When the operation of a machine involves repetitive work-cycles, it is possible (and useful) to identify a model with time varying model parameters. The time factor is incorporated into the system. This addition to the parameters (i.e., time dependent components) essentially allows the system increased degrees of freedom and therefore better solutions are expected. This system was also solved using all eight sets of measurements simultaneously, but ensuring that their time of collection was appropriately input into the model. A total elapsed time of 9.56 hours was used in the program with all the data from the experiments. The optimal solution provided a maximum error of 0.00027 inches for the entire work-cycle suggesting rather good performance across a 10 hour operating interval. If the work-cycles of the machine are not repetitive, and if jobs arrive randomly to the machine, the use of time dependencies will probably not provide an accurate characterization of the error variation. In such circumstances, it might be possible to substitute time dependency with temperature

dependency. This would require the identification of a few strategic locations on the machine structure where the temperatures are strongly correlated to the model parameter values. A methodology similar to the one described here can then be used to identify an adequate error model for compensation.

This research develops a new method for the estimation of the error model parameters of a machine tool using a Chebychev norm. This type of analysis forces strict bounds on the errors produced by the machine and can be used as a performance measure for the effectiveness of software compensation schemes on machine tools. As with any scheme, such a parameter estimation technique is not a universally superior technique. Its appropriateness for a particular situation is a matter of judgement. A few suggestions are relevant here. This technique must only be used when the maximum absolute error is the most relevant performance measure. Further, the maximum absolute error is very susceptible to outliers in the data. Therefore, such a technique must only be used when measurement errors are tightly bounded and measurement noise is low. Given proper experimental design and good instrumentation, such a technique can be gainfully utilized in machine tool metrology technologies and practice.

References

- [1] Yuan J. X. Ni J. Chen, J. S. and S. M. Wu. Real-Time Compensation for Time-Variant Volumetric Errors on a Machining Center. *Journal of Engineering for Industry*, 115:472-479, 1993.
- [2] Vasek Chvatal. *Linear Programming*. Freeman and Company, 1983.
- [3] A. Donmez. *A General Methodology for Machine Tool Accuracy Enhancement-Theory, Application, and Implementation*. PhD thesis, Purdue University, West Lafayette, Indiana, 1985.
- [4] R. J. Hocken. Three Dimensional Metrology. *CIRP Annals*, 26:403-408, 1977.
- [5] R. J. Hocken and the Machine-Tool Task Force. Technology of Machine-tools, Vol. 5: Machine Tool Accuracy. Technical report, Lawrence Livermore Laboratory, University of California, Livermore, California, 1980.
- [6] P. S. Huang and J. Ni. On-Line Error Compensation of Coordinate Measuring Machine. *International Journal of Machine-Tools and Manufacturing*, 35(5):725-738, 1995.
- [7] V. S. B. Kiridena and P. M. Ferreira. Computational Approaches to Compensating Quasistatic Errors of Three-Axis Machining Centers. *International Journal of Machine-Tools and Manufacturing*, 34(1), 1994.
- [8] V. S. B. Kiridena and P. M. Ferreira. Kinematic Modeling of Quasistatic Errors of Three-Axis Machining Centers. *International Journal of Machine-Tools and Manufacturing*, 34(1), 1994.
- [9] V. S. B. Kiridena and P. M. Ferreira. Mapping the Effects of Positioning Errors on the Volumetric Accuracy of Five-axis CNC Machine Tools. *International Journal of Machine-Tools and Manufacturing*, 34(1), 1994.
- [10] V. S. B. Kiridena and P. M. Ferreira. Parameter Estimation and Model Verification of 1st Order Quasistatic Error Models for Three-Axis Machining Centers. *International Journal of Machine-Tools and Manufacturing*, 34(1), 1994.
- [11] Yuan J. Lo, C. H. and J. Ni. An Application of Real-Time Error Compensation on a Turning Center. *International Journal of Machine-Tools and Manufacturing*, 35(12):1669-1682, 1995.
- [12] R. McClure, M. Weck, and G. Petuelli. Thermally Induced Error in Technology of Machine-tools. Technical report, Machine Tool Task Force, Vol. 5: Machine-tool Accuracy, Lawrence Livermore Laboratory, University of California, Livermore, California, 1980.

A 5-Degree-of-Freedom Laser System for Rapid Machine Volumetric Error Measurement and Compensation

Quanhe Ma, Xiaohua Chu, Kam Lau
Automated Precision, Inc.
Gaithersburg, MD 20879

Abstract

Due to increasing global competition, and the demand for improved product quality, the acceptable size and form tolerances for machined parts are continuously tightening. With the advancement of microprocessor control capabilities and software techniques, recent years have seen dramatic refinements in machine error characterization and compensation methods. In the production environment it is very important, sometimes even critical, to minimize measurement time on the machine, so that it is affordable for machine owners to conduct the measurements.

There are two conventional ways to measure volumetric errors on a multi-axis machining center. One method is using a laser interferometer. For a 3-axis machine, there are 21 independent error components: three linear errors and three rotational errors for each axis, and three squareness errors between the axes. The alignments for the measurements are very time consuming and tedious, because standard laser interferometer systems measure each error components separately. It normally takes a week to measure and characterize a 3-axis machine. The other approach employs distance measurement methods, such as diagonal measurements (via laser interferometer) or telescoping (LVDT- or laser-based) ballbars, to measure the volumetric errors. While faster, this method will provide data for overall volumetric errors, but cannot accurately identify the sources of the errors.

This paper introduces a new 5-degree of freedom (5-D) laser system. The 5-D laser system combines angular, straightness and linear displacement measurements in a single unit. Specifically, this unique device allows one to perform a dual-axis angle, dual-axis straightness and a linear displacement measurement in a single pass with a single set-up, eliminating the need for special and separate optics for straightness and angular measurements. The new 5-D laser dramatically reduces characterization time, e.g. to less than a day from over a week for a 3-axis CNC machining center.

The accuracy of the 5-D laser is analyzed and the test results comparing with a laser interferometer are presented in the paper.

The 5-D laser system has been used on many different CNC machining centers. Some measurement results are presented in the text, along with a volumetric model based upon measurements taken along the three orthogonal axes.

Key words: machine tool metrology, laser measuring system, machine error characterization and compensation

LASER MEASUREMENT INSTRUMENT FOR FAST CALIBRATION OF MACHINE TOOLS

P. S. Huang and Y. Li

Department of Mechanical Engineering
SUNY Stony Brook
Stony Brook, NY 11794-2300

1. Introduction

Calibration of machine tools is an essential operation in machine tool industry. It is routinely performed to keep machines in their peak performance and therefore to produce quality products. Some important purposes of calibration can be summarized as follows: 1) Error mapping of CNC machines for error compensation, 2) acceptance test of newly acquired machine tools, 3) periodic calibration for optimized performance of machine tools, 4) troubleshooting, and 5) demonstration of quality to potential customers.

One primary requirement for machine tool calibration is that the calibration must be performed as quickly as possible. Time-consuming calibration is costly due to the machine tool downtime as well as the expensive skillful labor required. Current industrial practice in machine tool calibration relies mostly on laser interferometers, such as HP5529A Dynamic Calibrator from Hewlett Packard and the laser calibration system from Renishaw. The latest models of these laser interferometers combine sophisticated optics and electronics with powerful software to simplify setup procedure and speed up data sampling and analysis. Relatively fast calibration have been achieved. However, the calibration speed is still limited because the basic measurement procedure did not change. With a single setup, still only one error component can be measured. Considering the fact that a multi-axis machine tool typically has tens of error components, the calibration is still a lengthy, time-consuming process.

To address the above problem, two multi-degree-of-freedom error measurement systems were developed to simultaneously measure up to five error components [Ni *et al.* 1992; Huang and Ni 1993], which significantly reduced the time of calibration. In another approach, the Telescoping Magnetic Ball Bar (TMBB) was developed [Bryan 1982] to quickly check the accuracy of a machine tool without

measuring each individual error component. The capability of TMBB was further extended by the development of the Laser Ball Bar (LBB) [Ziegert and Mize 1993] which has much longer range of motion than TMBB. LBB can also provide the measurement of individual error components [Ziegert 1994], in addition to the quick check of machine tool accuracy.

In this research, we follow the first approach and propose a new laser instrument that measures 5-DOF error components simultaneously when used with either the HP or Renishaw laser calibration system. These error components include linear distance, horizontal straightness, vertical straightness, pitch and yaw errors. The instrument is designed as a compact device to be used as an accessory to the HP or Renishaw calibration system to provide the capability of fast calibration of machine tools. The measurement principle and experimental results of the instrument are presented in the following sections.

2. Principle of Measurement

Figure 1 shows the optical layout of the 5-DOF error measurement system. Here the linear distance error is measured by using HP5529A's standard linear distance optics which consists of the He-Ne laser head, a linear interferometer (LI) and two linear retroreflectors (RR1, RR2). The laser beam is split into two beams by the linear interferometer, one reflected from the stationary retroreflector RR1 as the reference beam and the other reflected from the moving retroreflector RR2 as the measurement beam. Two non-polarizing beam splitters BS1 and BS2 and an interference filter are inserted into this standard linear distance measurement optics. BS2 with a beam splitting ratio of approximately 1:1 is used to extract half of the beam reflected from RR2 for the measurement of the other four error components.

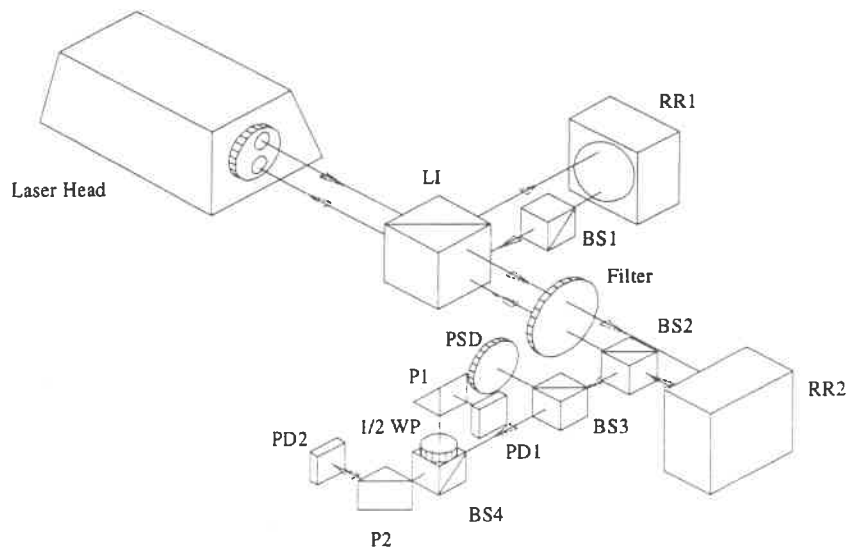


Figure 1 Optical layout of the instrument for 5-DOF error measurement.

BS1 is exactly the same type of beam splitter as BS2 and is inserted into the reference beam to balance the intensities of the reference and measurement beams for linear distance measurement. The band pass interference filter (peak wavelength 632.8 nm) is included to prevent the ambient light from reaching the photo sensors in the instrument and causing measurement errors.

The beam reflected from beam splitter BS2 is split again at beam splitter BS3. The reflected beam is directed to a two-dimensional position-sensing detector (PSD) for the measurement of horizontal (X axis) and vertical (Y axis) straightness errors. Following are the equations for calculating the straightness errors from the PSD signals:

$$S_x = K_1(x_1 - x_2)/(x_1 + x_2) \quad (1)$$

$$S_y = K_2(y_1 - y_2)/(y_1 + y_2) \quad (2)$$

where x_1 , x_2 , y_1 , and y_2 are the four output signals of the PSD. Note $(x_1 + x_2)$ and $(y_1 + y_2)$ are both measures of the intensity of the incident laser beam.

The transmitted beam at BS3 is further split at BS4 into two beams, one for pitch measurement and the other for yaw measurement. Pitch and yaw measurements are done by using a new angle measurement principle which is a variation of the previously developed angle measurement principle based on the Internal-Reflection Effect [Huang *et al.* 1992, 1994] with a simplified one-prism optical design. The pitch measurement unit consists of a right-angle prism and photodiode pair P1 and PD1 and the yaw measurement unit consists of another

right-angle prism and photodiode pair P2 and PD2. Following are the equations used to calculate pitch and yaw:

$$\theta_p = K_3(y_1 + y_2)/I_1 \quad (3)$$

$$\theta_y = K_4(y_1 + y_2)/I_2 \quad (4)$$

where I_1 and I_2 are the output signals of the photodiodes PD1 and PD2 respectively. Here pitch and yaw have been related to the inverse of the reflectance at right-angle prisms P1 and P2 respectively (note again that $y_1 + y_2$ is a measure of the intensity of the incident beam). We found that by so doing, the linearity is much improved in a certain angle range.

For the pitch measurement unit, a half-wavelength plate is also included to rotate the polarization of the laser beam so that the beam is reflected at P1 as a *p*-polarized beam, the same as that at P2 for the yaw measurement. Use of *p*-polarized beams is important in this application because the reflectance of a *p*-polarized beam is more sensitive to angular changes than that for an *s*-polarized beam.

Figure 2 shows the overall view of the measurement setup for simultaneous measurement of 5-DOF error components. All the optical components for measuring the additional straightness, pitch, and yaw errors are packaged into a small box, the 4-DOF Optical Head, that moves with the moving object. All the fixtures used in the setup are standard fixtures that come with the original laser calibration system.

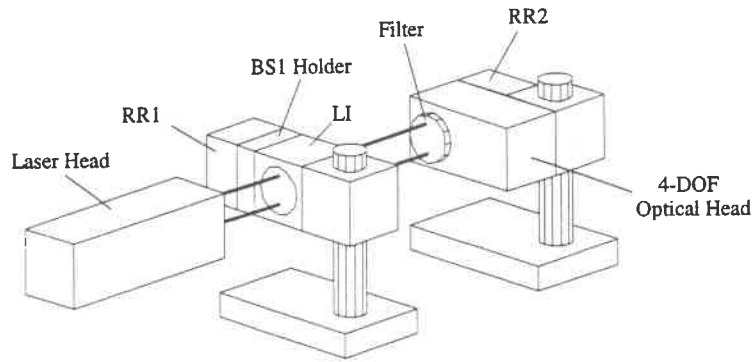


Figure 2 Overall view of the instrument for 5-DOF error measurement

3. Experiments

All the error measurement units except the linear distance measurement unit have to be carefully calibrated. In this research, we used a digital micrometer for the calibration of straightness error measurement and a precision angle measuring sensor [Huang and Xu 1995] for the calibration of pitch and yaw measurements.

Figure 3 shows the calibration result of the vertical straightness measurement unit. It is the average of six repeated calibrations. The calibration range is $\pm 200 \mu\text{m}$. From linear regression analysis, we found the maximum nonlinearity error to be approximately 1%. Better linearity can be obtained by using a better but usually more expensive PSD.

Figure 4 shows the calibration result for the yaw measurement unit. Again it is the average of six repeated calibrations. The angle range covered is $\pm 100 \text{ arcsec}$. The maximum nonlinearity error of yaw measurement is approximately 1.3%. Calibration of horizontal straightness and pitch measurements showed similar results.

Figures 5 and 6 show the stability of the vertical straightness and yaw measurements in 20 minutes. The short term fluctuation of the vertical straightness measurement is small ($< 0.5 \mu\text{m}$), but there is a clear trend of drift. The problem could be caused by thermally induced errors. For yaw measurement, the short term variation is approximately 2 arcsec peak-to-peak. Investigations revealed that the instability of the yaw measurement was mostly caused by the noise in the $(y_1 + y_2)$ signal. Measures are being taken to remedy the problem. Stability tests of the horizontal straightness and pitch error measurements showed similar results.

The instrument is currently being modified and further evaluated. Once this is done, it will be used to

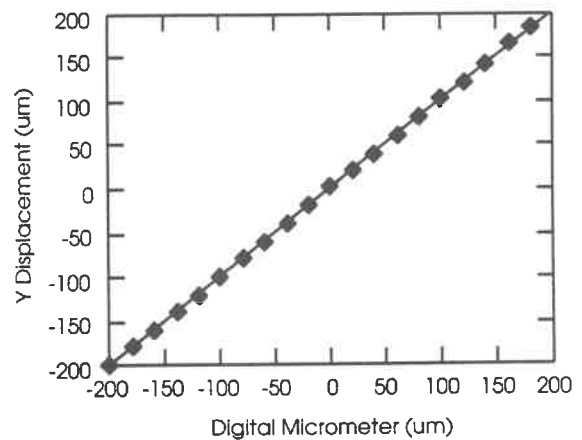


Figure 3 Calibration result for the vertical straightness measurement unit.

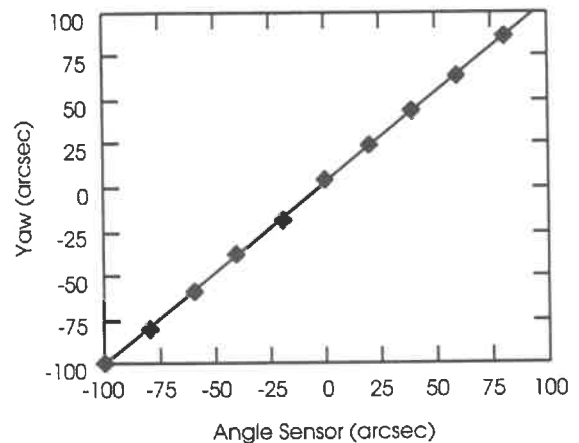


Figure 4 Calibration result for the yaw measurement unit.

calibrate the geometric errors of a CNC milling machine. The results will be reported shortly.

4. Summary

We have developed a new laser instrument for the simultaneous measurement of 5-DOF geometric errors of linear motion. The instrument has been designed as an accessory to the industry standard laser calibration system HP5529A from Hewlett-Packard. The instrument should also work with the Renishaw calibration system. When combined with the linear distance measurement optics of the above calibration systems, it provides the capability of simultaneous measurement of five error components, linear distance, horizontal straightness, vertical straightness, pitch, and yaw errors. The fast setup and measurement nature of the new instrument should significantly improve the efficiency of the machine tool calibration process.

References

- [1] Bryan, J. B., "A Simple Method for Testing Measuring Machines and Machine Tools, Part I: Principles and Applications," *Prec. Eng.* 4(2), 61-69 (1982).
- [2] Huang, P. S., Kiyono, S. and Kamada, O., "Angle measurement based on the internal-reflection: a new method," *Applied Optics*, 31, 6047-6055 (1992).
- [3] Huang, P. S. and Ni, J., "Angle measurement based on the internal-reflection effect and its application in the measurement of geometric errors of machine tools." *Proc. ASPE Annual Meeting*, 350-353 (1993).
- [4] Huang, P. S. and Xu, X., "Optical Probe for Surface Profiling of Aspherical Mirrors," *Proc. ASPE Annual meeting*, 112-115 (1995).
- [5] Ni, J., Huang, P. S., and Wu, S. M., "A multi-degree-of-freedom measuring system for CMM geometric errors," *ASME J. Eng. for Industry*, 114, 362-369 (1992).
- [6] Ziegert, John C. and Mize, Christopher D., "Design and Testing of a Laser Ball Bar for Rapid Machine Tool Metrology," *Proc. ASPE Annual Meeting*, 86-89 (1993).
- [7] Ziegert, John C., "Measurement of Machine Tool Parametric Errors," *Proc. ASPE Annual Meeting*, 76-79 (1994).

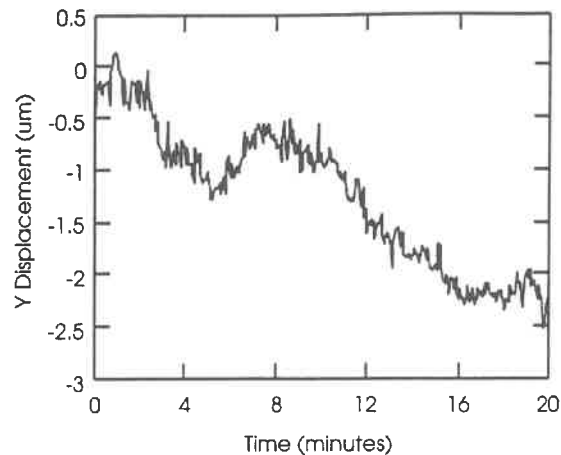


Figure 5 Stability of vertical straightness measurement.

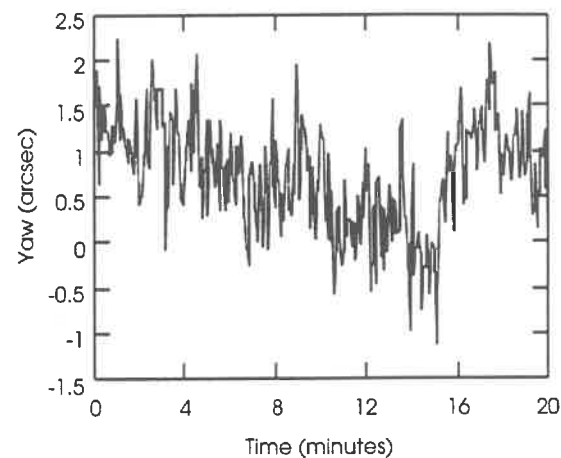


Figure 6 Stability of yaw measurement.

Practical Thermal Error Correction for CNC Machine Tools

J.P. Allen, S.R. Postlethwaite, D.G. Ford

The University of Huddersfield

Huddersfield, England.

Introduction

Internal heat generation and changes in environmental temperature inevitably produce changes in the temperature of machine tool structures. The resultant thermal distortion is widely recognised as a major cause of component inaccuracy. Design evolution has given improved thermal stability, through the development of more efficient spindle and axis drive systems and the use of cooling. Unfortunately the relentless drive for improved accuracy combined with increased cutting speeds and feed rates ensures that thermal errors remain a significant problem. Indeed it has been stated that in some cases thermal errors can account for up to 70% of the total machining error.

The aim of this research was to develop techniques for improving the thermal behaviour of general CNC machine tools beyond what is cost effectively achievable by good mechanical design. It was decided that compensation, where errors are corrected by adjusting the position of the tool or workpiece via existing machine axes, was the most promising basis for these general thermal error reduction techniques.

Thermal errors are dependent on the temperature distribution experienced by the machine tool structure. The operating conditions of most machine tools means that the heat generated internally varies significantly as the machine is used. This, combined with environmental temperature variation, can give rise to complex transient temperature distributions, making the resultant thermal errors difficult to quantify or predict. The principal problem facing the designer of a thermal error compensation system is determining how much correction is required at a particular instant.

The methods used by previous researchers to determine compensation values can be divided into two broad categories.

Direct Measurement: This usually takes the form of probing a reference face or an artefact mounted permanently on the machine. Systems based on direct measurement can give accurate results and the simpler ones are very effective at tracking and correcting slowly changing errors. Problems occur with more complex systems which require specialist hardware that must be specially tailored to fit individual machines and which can intrude on the working volume of the machine. Significant problems are also encountered when attempting to track rapidly changing errors, as each measurement operation requires a break in machining. Rapidly changing errors require frequent measurement operations and the result can be unacceptable time delays.

In-direct Measurement: Generally these systems calculate thermal errors from spot temperature measurements made on the surface of the structure. These systems can also give accurate results and have the advantage that errors can be tracked and corrected in the background. Additional hardware is fairly simple and does not intrude on the machine's working volume. The principle difficulty lies in determining the relationship between measured temperatures and compensation values. The majority of researchers have used large amounts of empirical data to train purely statistical models which relate temperature directly to error. They have not had access to a detailed picture of the temperature distribution produced by the machine and so they have tended to position temperature sensors intuitively. The result is that compensation can be very time consuming to implement. The models are also entirely statistical and so are not readily transferable even between machines of the same type.

Our Research Philosophy

This research project has attempted to address the weaknesses of previous research which have prevented thermal compensation being adopted widely. We have looked at developing techniques which can be implemented quickly, cost effectively and which are applicable to a wide range of machines.

Our research can be divided into two parts:

Part 1: The development of methods for the assessment of machine tool thermal behaviour.

Part 2: The development of general techniques for determining compensation values.

Methods for Assessing Machine Tool Thermal Behaviour (Part 1)

The purpose of this aspect of the research is to simplify the problem of applying compensation to a machine by identifying which aspects of its thermal behaviour produce the largest machining errors.

Machine Assessment Using Thermal Imaging: A machine tool can be considered as being made up of a number of discrete structural elements. Under normal operation some structural elements will undergo greater temperature changes and so are likely to produce greater thermal errors. A sequence of thermal images, recorded whilst the machine runs through a duty cycle, gives a detailed picture of temperature distribution and allows significant structural elements to be identified easily. An image of a Cincinnati vertical spindle NC milling machine is included here to illustrate the strength of thermal imaging in this application.



As expected the image shows a significant temperature rise of the headslide due to heat generated by the main spindle drive and bearings.

What is more interesting is a major, unexpected heat source within the column of the machine. Further investigation revealed that, because this is a relatively old machine, the spindle is driven via a gearbox mounted in the column. The thermal camera highlights the unexpected heat source immediately, giving not only its location but also the form of the temperature distribution it produces. It would be difficult to detect the presence of the gearbox using spot temperature measurements, let alone to obtain the same sort of detailed picture of temperature distribution. Without this

knowledge it could be difficult to model the thermal behaviour of this machine.

Machine Assessment Using Measurement of Error: In addition to the thermal imaging tests equipment is set-up to measure thermal errors directly. These measurements are used in conjunction with the thermal images to assist with the machine analysis. The equipment we normally use is non-contact sensors and a spindle mounted mandrel for measuring drift and a laser interferometer for measuring positioning change.

Error measurements allow us to assess the relative magnitude of different thermal errors. We also use the measurements in conjunction with the thermal image sequence to attribute thermal errors to particular structural elements.

Assessing the Effect of Thermal Distortion on Workpiece Accuracy: The aim of this part of the work was to develop an understanding of the ways in which thermal deformation of the structure

affects workpiece accuracy. We were able to define two general categories of thermal error which has proved useful for further simplifying the problem and for helping to select the most suitable compensation technique. These two categories of thermal error are:

Position Independent Thermal Errors (PITE): A change in the machine offsets. PITE change with temperature but not with position.

The effect of PITE on component accuracy is strongly dependent on the rate of change of PITE relative to the time taken to produce a component. Even large changes in PITE may result in little error if there is no significant change during the time taken to machine a component.

Slowly changing PITE can, however, be particularly significant if the component is rotated as part of the manufacturing process and the distance between cuts made before and after rotation is critical. Any offset that exists between the tool and workpiece will be reproduced on the component and if the component is reversed then the total error will be double the offset. If the component is not moved then, although the offset will appear on all the machined surfaces, the distances between surfaces should be accurate.

Position Dependent Thermal Errors (PDTE): A change in the linear positioning of the machine. PDTE change with temperature and with position. PDTE will produce errors if the change in linear positioning of the machine does not match the change in linear positioning required by the thermal expansion of the component.

Thermal Error Compensation Techniques (Part 2)

This section outlines some of the compensation techniques we have developed. Each technique is illustrated by a case study.

Novel In-Direct Measurement Based Compensation: A 3-axis CNC machining centre was producing rapidly changing PITE in the Y direction. An initial investigation revealed that the errors were produced mainly by thermal distortion of the headslide, which is heated by the main spindle bearings. The errors were changing too rapidly to be tracked efficiently by direct measurement so a novel technique for determining compensation values from spot temperature measurements was implemented.

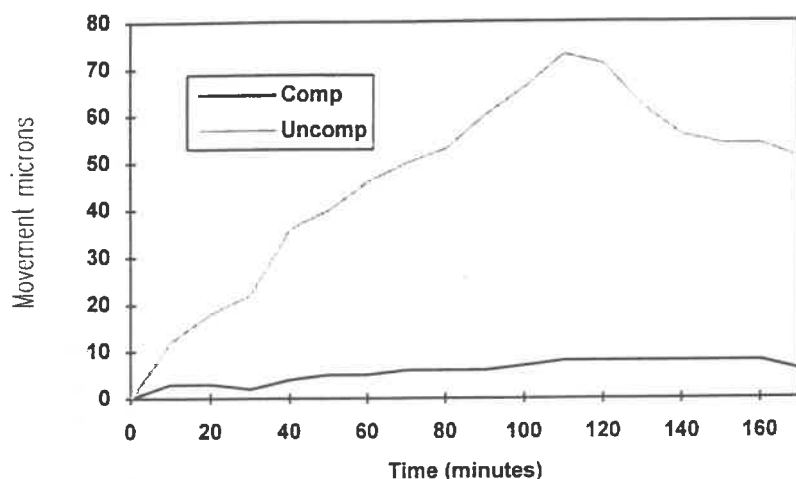
We have taken a more analytical approach to in-direct measurement based compensation than that taken by previous researchers. Rather than use a single statistical model to relate temperature measurements directly to error we use two distinct models, the temperature model and the distortion model, to calculate compensation values.

The temperature model calculates temperature distribution from spot temperature measurements. We have developed software which will analyse a sequence of thermal images to automatically choose optimum sensor locations and to optimise the calculation of temperature distribution. Normally separate models are built for each significant structural element. A simple heating-cooling test, with the camera focused on the structural element, recording images at intervals, normally provides enough data to build an accurate model.

The distortion model is used to calculate distortion from temperature distribution. It is usually analytically based but may feature some statistical tuning.

We have achieved success with this approach. Its main advantage over previous systems is that we are able to calculate thermal errors accurately using models built on data acquired during relatively short tests.

The graph below shows the results of compensation based on temperature measurements of the machine headslide. These results were obtained from a heating and cooling test that lasted approximately 3 hours.



The system works well with the compensation holding errors within a ten μm band. The rate of change of error has been greatly reduced which should have a significant effect on workpiece accuracy. The models used are able to deal equally well with the temperature distributions produced by both heating and cooling.

Implementation of Intermittent Re-datuming on a CNC Planing Machine: A CNC planing machine was being used to produce symmetrical parts. During planing the parts were supported in vees on two ground diameters, produced in a previous operation. The parts were planed on one side, rotated through 180° and planed on the other side. The machine user was having difficulty achieving the tolerance for the spacing between the faces produced before and after rotation.

An initial survey was conducted which found that the main thermal effect was a slowly changing PITE in the Z direction, tool moving normal to the planed surface. Although the total change in PITE over an 8 hour period could be around $80\mu\text{m}$ the change over the time required to produce a single component would only have been around $10\text{--}15\mu\text{m}$. Yet the user was experiencing errors of up to $100\mu\text{m}$.

A detailed study of the manufacturing process pin-pointed the problem. The error was produced as a result of the way the machine was referenced to the workpiece. Although the machine was under CNC control the referencing process was still carried out manually. Referencing was achieved by bringing the tool to a point $25\mu\text{m}$ away from a ground diameter. Whilst this worked well it was only being carried out for the first workpiece in a batch, the user was relying on the stability of the machine for producing subsequent components. It may take several days to produce all the components in a batch during which time the machine could move significantly. As the workpiece is supported on a previously ground diameter any shift in the reference position produces an error in the workpiece. Reversing the component effectively doubles any error.

The PITE changes slowly and there is a single dominant direction of movement. Consequently it could be tracked by probing at quite large time intervals, which should not cause any significant machining delays. In this particular case probing was not available on the machine but we recommended that referencing be carried out for each new component and repeated immediately after the component was rotated. This has effectively solved the problem and the user has reported no problems achieving tolerances.

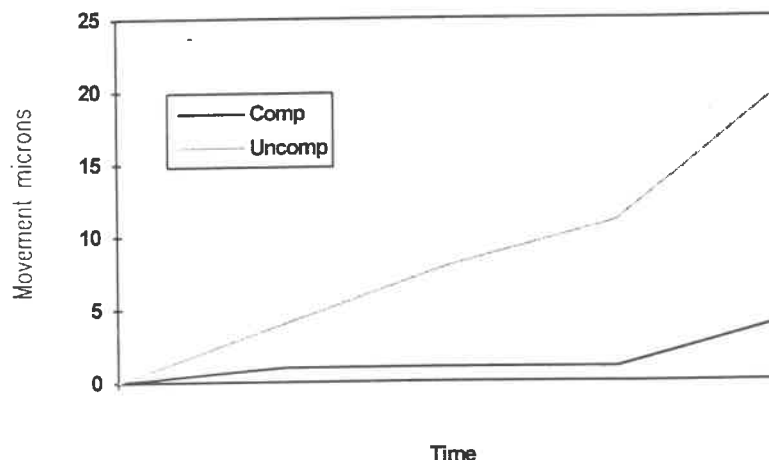
Although this is a fairly simple case it illustrates the strength of gaining a thorough understanding of the machine's behaviour and using this to decide on the best solution. In this case the solution was implemented at virtually no cost and resulted in reduced rework and scrap.

Use of a Thermally Stable Bar to Measure PDTE: Errors produced by change in linear positioning (PDTE) were experienced on a 3-axis CNC milling machine. An initial investigation on a single axis was carried out. The machine was fitted with rotary encoders mounted on the servo motors. The ballscrews were axially clamped at one end with the other end being free to slide. As a consequence the ballscrew itself is a major source of PDTE.

The first stage was to fit a linear scale to take the ballscrew out of the positioning loop. This reduced the PDTE from about 50µm change over the whole axis travel to 20µm, for similar duty cycles. In addition the reversal error in the axis was reduced to virtually zero. If free ended ballscrews are fitted to a machine then we would recommend linear scales as a prerequisite to achieving thermal stability.

The next stage was to track the remaining PDTE. To do this a thermally stable INVAR bar was mounted alongside the scale. The bar was clamped at one end and supported along its length on mounts that allowed axial movement. A displacement transducer measured relative movement between the free end of the bar and the scale. If we assume that the PDTE changes linearly then the thermally stable bar gives a direct measure of the expansion of the scale.

The graph below gives the results of a test. The data shows the change in positioning error over the whole axis length. The changes were produced by a combination of environmental temperature change and heat produced by cycling the axis at 2 m/min.



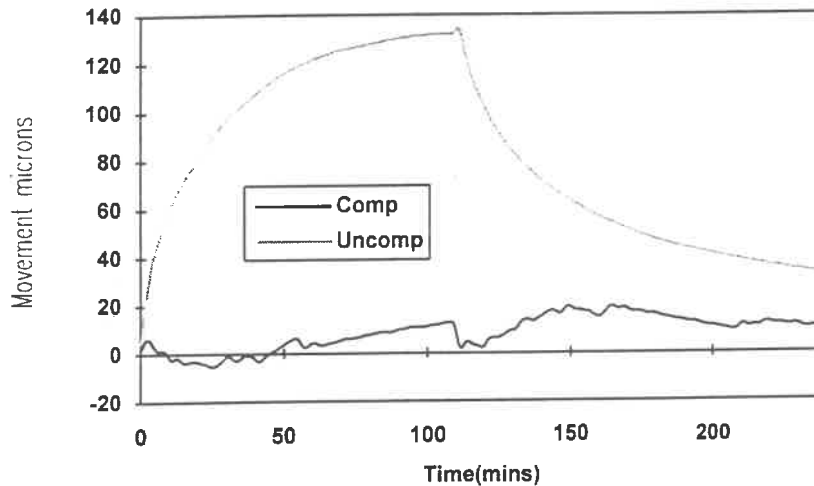
In this case the thermally stable bar gives a good measure of the error. Using a combination of a linear scale and the thermally stable bar allows the change in linear positioning to be reduced to less than 10% of its original value.

Direct Measurement Based Compensation Applied to a CNC Lathe: The final case concerns the thermal errors produced by a large CNC lathe. The initial investigation revealed rapidly changing PDTE in the Z direction (along the spindle axis). These errors were produced by thermal expansion of the spindle shaft itself. The spindle had a low thermal mass and was heated directly by the main spindle bearings. As a consequence it heats and cools very rapidly. The spindle is supported in two sets of bearings. One provides axial location while the other is free, to prevent the bearings being overloaded by spindle shaft expansion. In this case the rear bearings were clamped causing the spindle to expand forwards out of the headslide, resulting in large machining errors.

Normally when faced with rapidly changing PDTE we would use temperature sensor based compensation to track the errors. In this case the most significant structural element was the spindle

shaft. The rotation of the spindle made accurate temperature measurements difficult to obtain. Fortunately the initial investigation had given the information necessary to develop a simple direct measurement based solution. We mounted a non-contact displacement sensor on the machine to measure relative movement between the headstock and the back of the chuck. The signal from the sensor was used to provide a compensation signal for the Z axis drive.

The graph below shows the Z axis thermal error, with and without compensation, for a heating and cooling duty cycle.



The system works well with the compensation holding errors within a 20 μ m band.

The technique will cope with errors produced by any relative movement between spindle and headstock.

Conclusions

- Techniques have been developed which use thermal imaging and error measurement to rapidly assess machine tool thermal error behaviour.
- A general philosophy of machine tool thermal error behaviour has been formulated. This can be used to help in the analysis and categorisation of individual machines.
- A range of practical techniques have been developed to address the specific thermal problems of individual machine tools.

Methodologies for the Evaluation of Systematic Form Deviations for Inside-Diameter Cylindrical Features and Their Relationship to Process Variables¹

R. P. Henke, R. M. Cassou, K. D. Summerhays, University of San Francisco, San Francisco, CA; J. M. Baldwin, Sandia National Laboratories, Livermore, CA; and C. W. Brown, AlliedSignal Federal Manufacturing & Technologies, Kansas City, MO

ABSTRACT

The ever-increasing demands for precision machined parts has created a renewed emphasis on better understanding of the relationships between process control variables and deviations from perfect geometric form. To meet this demand, it is important to have high-precision, high-density metrology data on part features which (in some sense) span the range of variability in form for the process under study. Clearly the problem is a complex one involving the machine tool itself, materials, fixtures, cutters, feeds, speeds, and many other factors. In the work described here we introduce a methodology which can serve to identify key analytical deviations from form and their dependence on several of the factors enumerated above.

Artifact parts containing inside-diameter cylinders of various diameters and depths were produced in aluminum in statistically significant quantities on a flexible manufacturing system. Blind, counter and through holes were included. Manufacturing methods included drilling, boring, and reaming operations in various combinations. Nearly 1400 of these cylinders were measured on a high-precision coordinate measuring machine adapted for high-density sampling. Sample points (about 1000 per cylinder, taken on a uniform grid) were fitted by least-squares as a linear combination of analytic terms.

In most instances the systematic deviations from perfect geometric form could be conveniently modelled by treating the form deviations according to a Fourier series in the angular dimension and a simple polynomial in the axial dimension. A basis set of specific cross-terms in the resultant product function has served to represent what are often taken as the classic deviations from standard form, eg. taper, lobing, axial curvature, etc. One additional readily recognizable basis shape, not manifested in any *one* cross-term, but in a set of correlated terms, has been identified as a result of this work. The relative importance of the form deviation components in explaining the variance over a range of individual features has been assessed. In many cases the majority of the variance observed from ideal part geometry can be ascribed to fewer than a half-dozen such terms. The nature of these terms and their ranking in accounting for the variance is strongly influenced by the specific manufacturing method employed. The expression for the form deviations was further simplified by a principal component analysis of the shapes that result from a given manufacturing technique. This reduced the basis set to the minimum number of functions necessary to capture the resulting variability in the set of features manufactured by this technique.

In the characterization of cylinders produced by a succession of operations such a center-drill/drill/bore/ream, the importance of each step in the sequence is considered. Certain generalities have emerged which should aid in establishing the domain of process variable control.

METHODS

The deviation of a manufactured cylinder from perfect form must be expressed quantitatively and in a manner which is much more concise and abstract than the dense set of points used to measure the cylinder. For this reason the deviation has been expressed as a weighted sum (linear combination) of basis functions, each of which can be pictured as a shape over the cylinder which is to be multiplied by a coefficient and then added to the existing cylinder in a radial direction. One functional form used is the set consisting of all terms such as $T_n(z)$, $T_n(z) \cos(m\theta)$, or $T_n(z) \sin(m\theta)$. Here $T_n(z)$ is the Chebyshev polynomial of the first kind and order n , where $n \geq 0$ and the axial coordinate z has been normalized such that $-1 \leq z \leq 1$. The Fourier order $m > 0$ and θ is the polar angle. If the maximum order in θ is M and in z is N then we can define a basis function vector, $b(z, \theta)$, which is to be evaluated at any point on the surface of the cylinder.

$$\mathbf{b}(z, \theta) = (T_0(z), T_0(z) \cos(\theta), T_0(z) \sin(\theta), T_0(z) \cos(2\theta), T_0(z) \sin(2\theta) \dots T_0(z) \cos(M\theta), T_0(z) \sin(M\theta) \dots \\ T_1(z), T_1(z) \cos(\theta), T_1(z) \sin(\theta), T_1(z) \cos(2\theta), T_1(z) \sin(2\theta) \dots T_1(z) \cos(M\theta), T_1(z) \sin(M\theta) \dots \\ \dots \\ T_N(z), T_N(z) \cos(\theta), T_N(z) \sin(\theta), T_N(z) \cos(2\theta), T_N(z) \sin(2\theta) \dots T_N(z) \cos(M\theta), T_N(z) \sin(M\theta))'$$

As can be seen $\mathbf{b}$ is an $(N+1)(2M+1)$ dimensional vector. Then the radius of the cylinder at the point z, θ is given by $r = R + \mathbf{b}' \mathbf{a}$, where R is the average or perfect form radius of the cylinder, $\mathbf{a}$ is the vector of function coefficients, and $\mathbf{b}' \mathbf{a}$ is the inner or "dot" product of the vectors $\mathbf{b}$ and $\mathbf{a}$. In practice $\mathbf{a}$ is determined by a least squares fit to at least $(N+1)(2M+1)$ measured points on the surface of the cylinder. The vector $\mathbf{a}$ represents the deviation from perfect form, so all of its components are assumed to be much smaller than R .

The representation given above is useful in several respects. 1) If we believe we have used orders in z and θ which are sufficiently high to represent the systematic component of the form error, we can determine $\mathbf{a}$ from relatively few points on the surface and then evaluate the radius of the actual "cylinder" at any point on its surface. How is it possible to know the required order in z and θ ? Densely measuring and then least squares fitting a number of surfaces which are typical of those produced by a given manufacturing technique allows a determination of the z and θ orders of the terms which are statistically significant. 2) Once we have the functional form it is possible to use experimental design techniques² to find the optimal locations on the cylinder surface to provide the best determination of $\mathbf{a}$ from a minimum number of points. This determination has been one of the main thrusts of our research. 3) Knowing which terms are most significant (which components of $\mathbf{a}$ are large) provides a much better conceptual picture of the character of the form deviations than is provided by mere data sets alone. This is because certain terms in the above expansion can be identified with commonly known and named forms of error. Table I gives a list of the descriptions of some of the terms.

Table I

Term	Description
$T_0(z)$	Small change in radius
$T_0(z) \cos(\theta)$	Small change in x location
$T_0(z) \sin(\theta)$	Small change in y location
$T_0(z) \cos(n\theta), T_0(z) \sin(n\theta)$	n lobes
$T_1(z)$	Taper
$T_1(z) \cos(\theta)$	Small change in direction cosine i
$T_1(z) \sin(\theta)$	Small change in direction cosine j
$T_2(z)$	Hourglass/barrel
$T_2(z) \cos(\theta), T_2(z) \sin(\theta)$	Banana

One of the serious drawbacks of the Chebyshev/Fourier representation is that it is often necessary to go to a fairly high order in z and/or θ to represent the structure present in the data. When this happens it is reasonable to include all terms up to the maximum order in the polynomial. This results in a large number of terms, many of which are not statistically significant. Many points are required to determine all of the coefficients and much of weight of the data goes toward determining parameters which are not significant. A better approach is to capture just the right linear combinations of the Chebyshev/Fourier functions. These composite functions provide a basis set which can

best fit the structure present in a representative sample of cylinders. The technique used to find the proper basis set is principal component analysis. It is designed to represent the largest possible fraction of the variance present in the cylinder set with the minimum number of basis functions.

We now consider the error associated with the deviation from perfect form given over a cylinder by the function $b'(z, \theta)$. The goal of this procedure is to find the minimum number of basis functions that represent the variance over a large set of surfaces. The procedure is somewhat complicated so the reader may proceed to the next paragraph if he or she is willing to take the steps for granted. We choose a large number of analysis points, uniformly spaced over the surface, and create one row of a matrix, B , for each point on the surface. This row will be $b'(z, \theta)$. The product $B a$ is a column vector, each element of which is the deviation at one point on the surface. It is assumed at this point that the orders in z and θ are large enough to capture all of the systematic shapes contributing to the surface form deviations. (In practice we use 9th order for each of them, which leads to 190 basis shapes.) Now imagine a number of densely measured cylinders which represent all of the possible variation for the manufacturing method in question. Least squares fits of each of these results in a set of a vectors. We construct the matrix A by using each of the a vectors as a column of A . The product $S = B A$ is a matrix whose columns are the vectors which give individual surfaces. The trace of $C = S S'$, divided by the number of points per generated cylinder surface and again divided by the number of surfaces, is the square of the rms error (the average variance) over all of the cylinders in the set. Our goal is to represent this average variance by as few basis set shapes as possible. Now $C = S S' = B A A' B' = B E V' E' B'$. The last step is accomplished by realizing that the product $A A'$ is the covariance matrix of the parameters in a times the number of cylinders in the set. We diagonalize this product using a standard single value decomposition. Since $A A'$ is symmetric its decomposition is also symmetric and can be expressed as $E V' E'$, where E is a matrix whose columns are the eigen vectors and V' is a diagonal matrix with the eigenvalues along the diagonal. Next we replace $B E$ with its product F' , which gives $C = F' V' F'$. Now to provide a normalized basis set (though not quite orthonormal), we normalize each of the columns of F' taken as a vector and multiply the corresponding diagonal element of V' by the square of its length. This leads to the normalized form $C = F'' V'' F'''$. Finally we simultaneously rearrange both the position of the diagonal elements of V'' and the columns of F'' to place the elements along the diagonal in V'' in descending order. With this last change we seek and we can write the modified expression as $C = F V F'$. At this point we don't worry about the fact that C is multiplied by the number of points/surface times the number of cylinder surfaces because we will be expressing the variance associated with each of the eigensurfaces as a fraction of the total.

We now have the product $C = F V F'$, which can be described as follows. The columns of F are the eigen normalized basis surfaces we seek, V is a diagonal matrix with descending size diagonal elements proportional to the variances associated with the corresponding column of F , and the trace of V is proportional (with the same proportionality constant) to the total variance over all of the surfaces. Hereafter we will assume that V is divided by its trace to make the diagonal elements the descending size fractional variances we seek.

RESULTS

Five different sequences of operations were employed in the manufacture of the 46 different cylindrical features on each of 30 nominally identical parts. The features included blind and through holes and counterbores with diameters ranging from 1/8" (3.2 mm) to 1" (25.4 mm) and lengths up to 0.997" (25.3 mm). Since our basis set as expressed in the vector b is normalized, it is possible to compare the form errors of cylinders of different diameters and lengths. The five different manufacturing sequences and their acronyms are: center drill/drill (CD), center drill/drill/ream (CDR), center drill/drill/bore/ream (CDBR), center drill/drill/plunge end mill (CDP), and center drill/drill/contour mill (CDM).

First we consider the fraction of the variance as explained by each of the Chebyshev/Fourier (polynomials) as a function of the manufacturing method. We are interested in knowing, as a function of the manufacturing sequence, which are important basis terms for expressing the form errors. All of the surfaces were least squares analyzed using the polynomial basis set. For each individual cylinder, taken over all 30 parts, the fractional variance explained by each of the basis functions was obtained. For all parts manufactured by a given sequence, the largest (worst) fractional error contributed by any of the holes was taken to be the significant requirement for inclusion of that term. Only cases where the fractional variance contribution was greater the 5% were registered to avoid overcomplicating

the result. The result of this study is shown in Fig. 1.

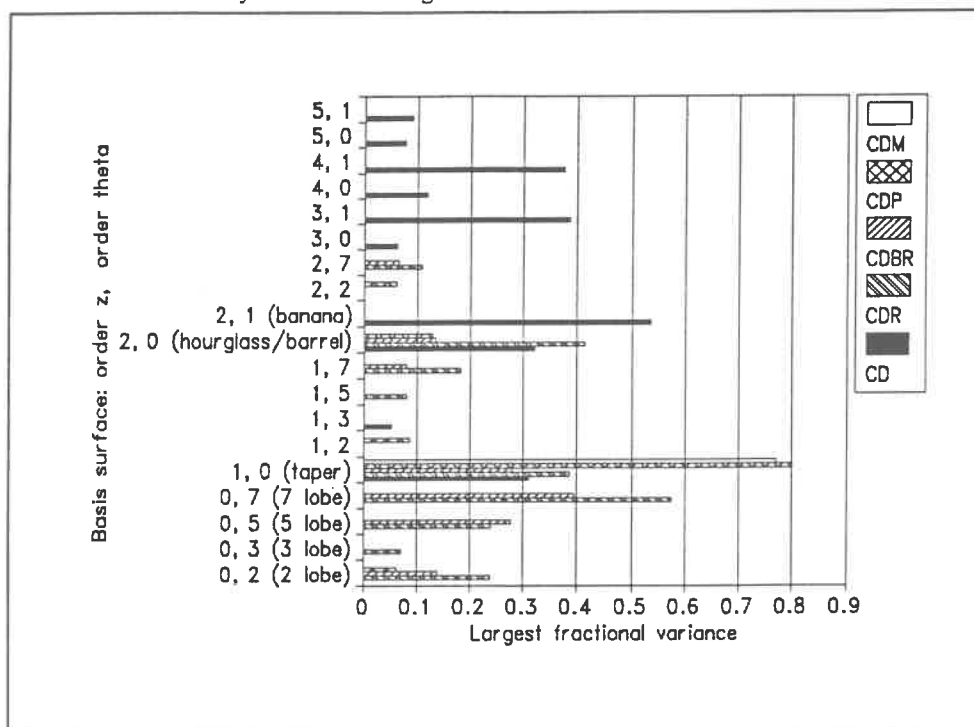


Figure 1. Analysis to find the necessary terms in a minimal polynomial basis set.

Next we consider the result of the principal component analysis. Each of the five manufacturing sequences was treated separately by including all cylinders made by that method. All terms through ninth order in both z and θ were allowed. The first result of this study is shown in Fig. 2. Presented is a comparison of the fraction of the variance which remains to be explained after n terms are included in the eigen basis set.

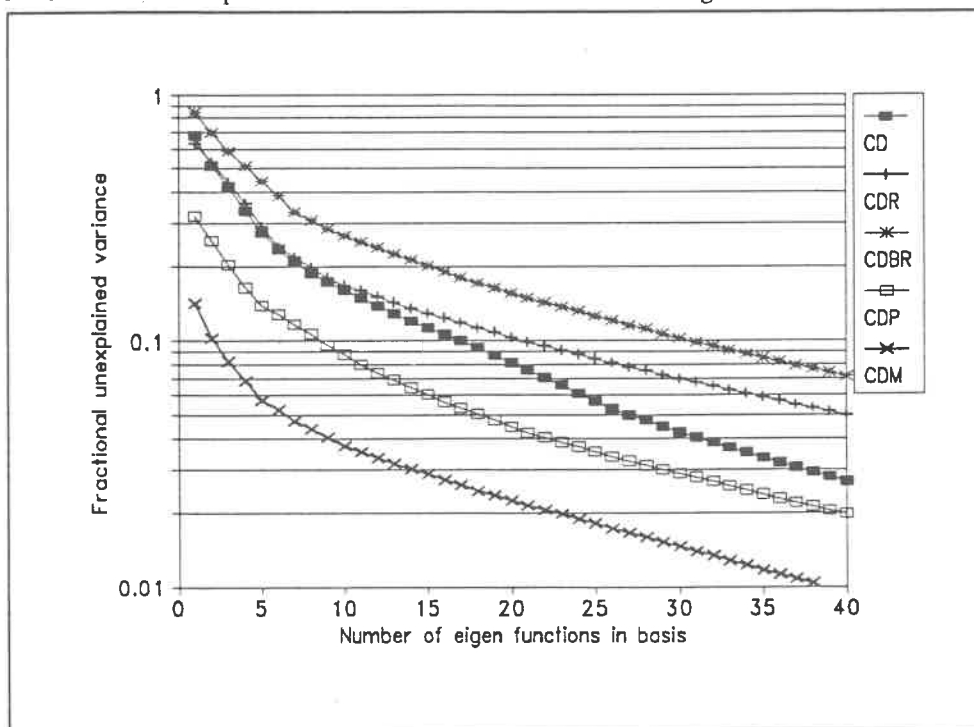


Figure 2. Fraction of the variance which cannot be represented by the basis set as a function of the number of eigen functions in the basis set.

Figs. 3-12 show two of the eigen shapes for each of the five manufacturing sequences employed. What is shown in each of these figures is six different views of the cylinder, each with a highly magnified rendition of the deviation from nominal form. The first four views are oblique side views. The arrow at the top of the cylinder indicates the direction of the origin in θ . The fifth view is a top view again with the arrow indicating where $\theta = 0$. The sixth view is a peeled out view of the inside of the cylinder (all of the cylinders were ID cylinders) with the cut being at the origin in θ . The near side is the top of the cylinder and the left side is $\theta = 0$.

DISCUSSION

Fig. 1 allows us to observe some generalities about the manufacture of ID cylinders. The most dominant behavior seen for all methods is taper. This is particularly true for holes generated with a final pass of a reamer. Another common error is hourglass. Both of these kinds of errors might be caused by lack of perfect concentricity of all of the elements of the machining operations. Reamed holes generated by an n -fluted reamer also show pronounced $n-1$ and $n+1$ lobed patterns. In Fig. 1 the CD method generates a number of predominant terms with both a higher order z and a θ dependence. It is not obvious, but can be seen from the eigen analysis, that these are mostly used to represent a spiral error structure.

The shapes shown in Figs. 3-12 may be said to be the essence of the errors from true form generated by the different manufacturing sequences. The dominant behavior for CD (Fig. 3) is a single-lobed right-handed spiral shape. It is designated as single-lobed because a cross section at any height is essentially a circle. Analysis of individual cylinders shows that the pitch of this spiral varies from about 1.3 to 4 times the pitch of the tool. CD eigen shape 2 is the other phase of the spiral and is not shown. The spiral shape was more predominant in smaller holes and was particularly pronounced in a 1/4" hole generated with a parabolic flute drill. It should be noted that the drills used have 2 flutes and produce a 1-lobe ($n-1$) pattern. The CD behavior that becomes more predominant for larger diameter holes is the eigen shape 3 shown in Fig. 4. It has a small banana shape combined with several different orders of z dependence.

The holes manufactured with the final pass being an n -flute reamer are dominated by $n-1$ and $n+1$ lobed patterns. They often also have a z dependence dominated by taper and/or hourglass. The CDR holes show both a nearly pure z dependence (Fig. 5), which is approximately an hourglass shape, and a lobed θ dependence (Fig. 6). The pattern shown in Fig. 6 is a 7-lobed pattern. The other phase of 7-lobe and both phases of a 5-lobe pattern also are present in the sequence of eigen shapes, but are not shown. Whereas the 7-lobe shape was more predominant in the CDR method, the CDBR method has a more predominant 5-lobe pattern (Fig. 7) but also has 7-lobe variation (Fig. 8). Notice that the CDBR holes are truer in the z direction than the CDR holes with the only z dependence in the first two eigen shapes being the slight taper in Fig. 7.

Holes generated by a final pass of an end mill are dominated by taper (Figs. 9, 11). They can also have very significant hourglass (Fig. 10). Fig. 10 also shows a 2-lobed spiral component. In the NC generated holes of the contour milling operation (Figs. 11, 12) the angles at which the tool entered and left the material are essentially the same from hole to hole. This can be seen as a slight hint in the shape dominated by taper (Fig. 11) and in a very conspicuous way in Fig. 12. The presence of this eigen shape 2 indicates this pattern varies from part to part.

Fig. 2 allows us to conclude that more complicated operations result in structures which vary from true form in a more complicated and detailed way. The operations which need few eigen shapes to explain the errors they generate can be said to have few degrees of freedom in the behavior of the tools as the holes are being generated. With each successive pass of a tool (which often is removing little material) the tool motion can couple with the existing pattern in the hole. That is, the previous pattern imprints itself on the following operation, which also produces its own contribution to the details of the shape of the hole. The holes generated by a mill, which is a reasonably rigid tool compared with either a drill or a reamer, essentially erase the effects of the preceding operation and might be considered to be a single-pass operation. This is particularly true of the contour milling operation, which removes much more material than the plunge milling operation. In the figure this is demonstrated by the fact that the milling operations require the fewest eigen shapes to explain the patterns they generate. The two-pass operation, CD, requires more shapes, the three-pass operation, CDR, requires even more, and the four-pass operation, CDBR, requires the most explanation.

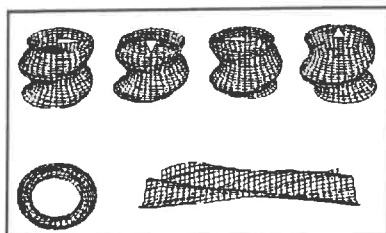


Figure 3. CD eigen shape 1.

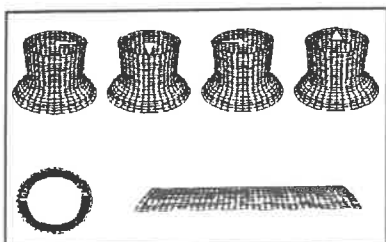


Figure 5. CDR eigen shape 1.

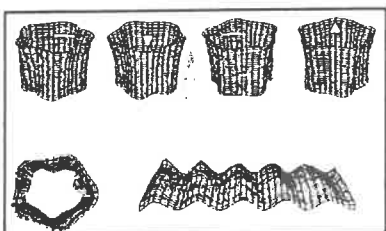


Figure 7. CDBR eigen shape 1.

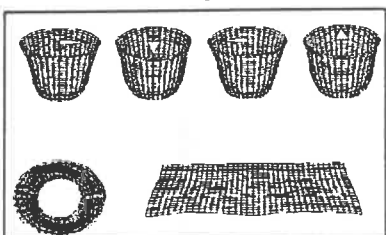


Figure 9. CDP eigen shape 1.

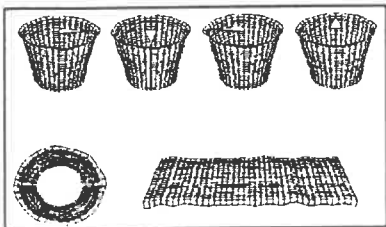


Figure 11. CDM eigen shape 1.

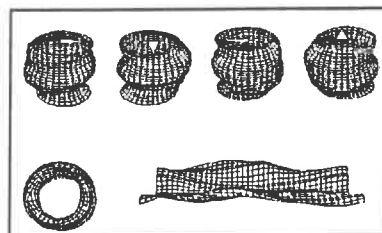


Figure 4. CD eigen shape 3.

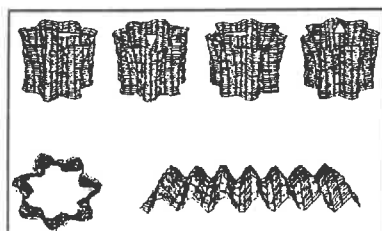


Figure 6. CDR eigen shape 2.

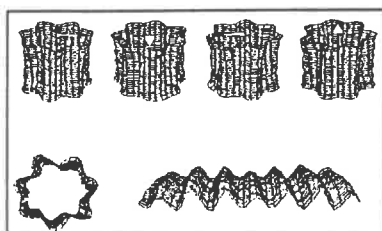


Figure 8. CDBR eigen shape 5.

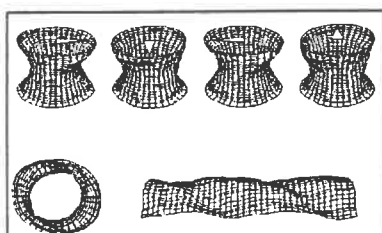


Figure 10. CDP eigen shape 2.

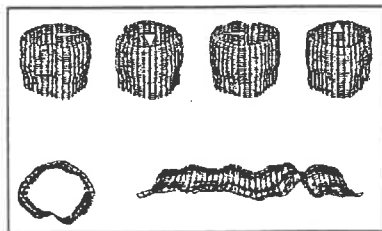


Figure 12. CDM eigen shape 2.

REFERENCES

¹Work supported in part by the Quality Assurance Program of the Institute for Manufacturing and Automation Research, Consortium for Advanced Manufacturing International, Bedford, Texas

²K.D. Summerhays, R.P. Henke, R.M. Cassou, J.M. Baldwin, and C.W. Brown (1995), *New Algorithms for the Evaluation of Discrete Point Measurement Data and for Sample Point Selection on Surfaces with Systematic Form Deviations*, Proc. 10th Annual Meeting of the American Society for Precision Engineering, Austin, TX, 388-391.

BEST FIT TRANSFORMATION AND UNCERTAINTY ANALYSIS IN COMPUTATIONAL METROLOGY

Hong-Tzong Yau
Department of Mechanical Engineering
National Chung Cheng University
Chia-Yi, Taiwan

Introduction

In computational metrology, the key process for measurement and tolerance evaluation is the fitting of geometric elements [1], or sometimes referred to as the best fit. The goal of the best fit is to create substitute geometry from discrete measurement points. Traditionally, various fitting algorithms and criteria are used to fit data points to different geometric elements such as line, circle, plane, and cylinder. Due to the divergence of methods and criteria used, the best fit results tend to be different from case to case. Efforts have been made in the U.S. and Europe to develop software testing systems that will evaluate CMM software errors by comparing the results against some known reference algorithms.

In the author's previous work [2], a unified best fit algorithm was proposed as a reference algorithm for software testing based on least-squares criterion (which is adopted by most CMM vendors). The proposed algorithm does not rely on approximation or linearization to reduce complexity or enhance speed, therefore can produce "exact" answers opposed to approximate solutions. More importantly, the method points out that the best fit can be viewed as a coordinate transformation between measurement data and nominal geometry. Therefore, best fit for any geometric feature can be unified by minimizing the following cost function:

$$F = \sum_{i=1}^m |Q_i - TS_i|^2 \quad (1)$$

where Q_i is a measured point, S_i is Q_i 's projected point on the nominal surface, and T is the coordinate transformation matrix. The proposed algorithm has proven to be robust and can generate accurate solutions verified by simulations and experiments [2].

Uncertainty Analysis

Even with a perfect algorithm, true surface deviations plus measurement errors and limited sampling data can lead to uncertainties in the best fit result [3]. Assessment of uncertainty becomes important as it reveals reliability of the best fit result and can guide the inspector to determine a proper CMM sampling plan. Since the best fit can be viewed as a coordinate transformation, uncertainties of the best fit result can be modeled as the possible variations of the transformation parameters. In a 3D best fit, this would involve 3 rotation parameters (roll, pitch,

and yaw) and 3 translation parameters (x, y, and z displacements). The variation of these parameters $\delta \tilde{t}$ can be related to surface form error $\tilde{\varepsilon}$ by a sensitivity matrix A [4].

$$\tilde{\varepsilon} = A \delta \tilde{t} \quad (2)$$

Assuming δt_i and ε_i are normally distributed random errors with zero means and let the sensitivity measure of δt_i to ε_i be defined as

$$S_{t_i} = \frac{\sigma_{t_i}}{s} \quad (3)$$

where σ_{t_i} and s are the standard deviations of δt_i and ε_i respectively, general sensitivity equation can be derived as [4]:

$$S_{t_i} = \sqrt{\sum_{j=1}^m \left[\left(A^T A \right)^{-1} A^T \right]_{i,j}^2} \quad (4)$$

This sensitivity measure relates uncertainties of the best fit transformation to surface form errors, measurement positions, and CMM sampling size.

The first attribute that affects the best fit uncertainties is the surface form error. Assuming sufficient measured points are sampled, based on equation (3), uncertainty is proportional to the standard deviation of the surface form error. In other words, the larger the form error, the more uncertain (and less reliable) the best fit result.

The second attribute considered is the sampling positions. Typically, given a part surface for measurement, uniform distribution is used to determine the sampling positions. In terms of reducing the uncertainty, "best" sampling positions (with a fixed number) can be obtained by minimizing the overall uncertainty values.

$$F = \sum_{i=1}^6 S_{t_i} \quad (5)$$

Since the sensitivity matrix A is a function of the sampled point $\vec{N}_j$ and its normal vector $\vec{n}_j$, the cost function becomes

$$F = F(A(\vec{N}_j, \vec{n}_j)) \quad j = 1, \dots, m \quad (6)$$

Yet, any surface can be parameterized as a function of the bivariate surface parameters (u, v). Therefore,

$$F = F(A(\vec{N}(u_j, v_j), \vec{n}(u_j, v_j))) \quad j = 1, \dots, m \quad (7)$$

The minimization of this function is a nonlinear optimization problem with 2m variables. In this work, conjugate gradient method is adopted to search for the optimal sampling positions given a fixed number of measurement points. Uniform parameter distribution is used to initialize the iterative search. Figure 1 shows the method applied to a complex die surface for finding its best measurement positions. Figure 2 shows the convergence of the cost function. It appears the improvement in lowering the uncertainty is definite, but the gain is not much (compared to the sampling size effect discussed later). The result seems to indicate that as long as measurement

points are well distributed, fine tuning of the sampling positions does not reduce the best fit uncertainties much.

The third attribute considered is the CMM sampling size. It is known that increasing CMM sampling size would reduce uncertainties. But can this effect be measured and quantified? Using the previous die surface as an example, sensitivity measures (proportional to uncertainty values) for the transformation parameters were calculated and plotted against the CMM sampling size (Figure 3). The sensitivity measures exhibit exponentially decreasing patterns as the sampling size increases. Based on the 2D circle experience [4], it is speculated that the sensitivity measure (uncertainty value as well) is in general inversely proportional to the squared root of the sampling size. A theoretical curve $S = K / \sqrt{m}$, where K is a constant, is used to fit the curves in Figure 3. K can be calculated by

$$K = \frac{\sum_i S_i / \sqrt{m_i}}{\sum_i 1 / \sqrt{m_i}} \quad (8)$$

Good matches were found between the theoretical curves and the original curves.

Once "calibrated", this curve can be used to characterize the relationship between uncertainty and CMM sampling size. For example, if the magnitude of the surface form error is increased by two, the CMM sampling size should increase by 4 to maintain the same uncertainty level.

The last but not the least important attribute is the geometric shape of the element. Consider a cone symmetric to the z-axis. The two parameters affecting the geometric shape are the height H and the radius R . If we vary H while holding R and plot the sensitivity measures of the x and y translation components, it results in interesting curves as in Figure 4. When $R = H$, uncertainty is at its lowest value, indicating when the cone angle is 90 degrees, the center position of the base circle is most certain. However, as H decreases, the uncertainty increases rapidly. On the other hand, when H is larger than R , the uncertainty value would increase somewhat and then saturates at a fixed value.

References

- [1] T. Hopp, 1993, "Computational Metrology", Proceedings of the 1993 International Forum on Dimensional Tolerancing and Metrology, CRTD-Vol. 27.
- [2] H.T. Yau, 1996, "A Unified Least-Squares Approach to the Evaluation of Geometric Errors Using Discrete Measurement Data", Int. J. of Machine Tools and Manufacture, in press.
- [3] S.D. Phillips, B. Borchardt, and G. Caskey, 1993, "Measurement Uncertainty Considerations for Coordinate Measuring Machines", NIST report 5170.
- [4] H.T. Yau, 1993, "Uncertainty Analysis of Circle and Arc Measurement Using Discrete Sample Points", Proceedings of the 1993 ASPE Annual Meeting.

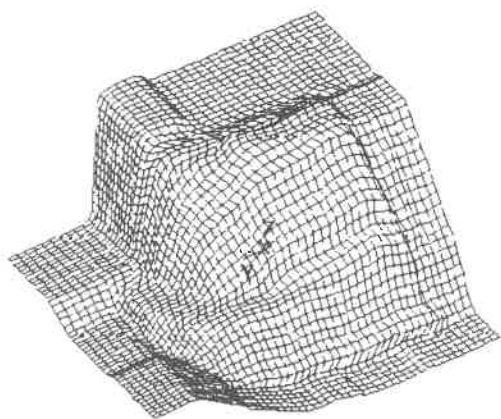


Figure 1. A complex die surface for uncertainty analysis.

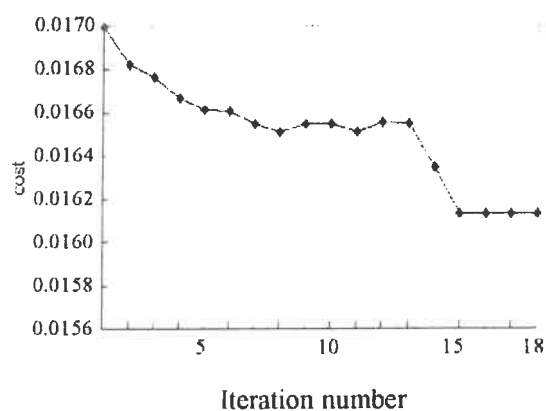


Figure 2. Convergence of the cost function.

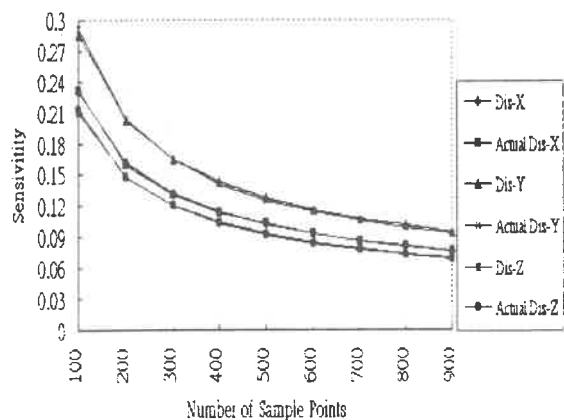


Figure 3. Theoretical and simulated sensitivity curves for the translation vector.

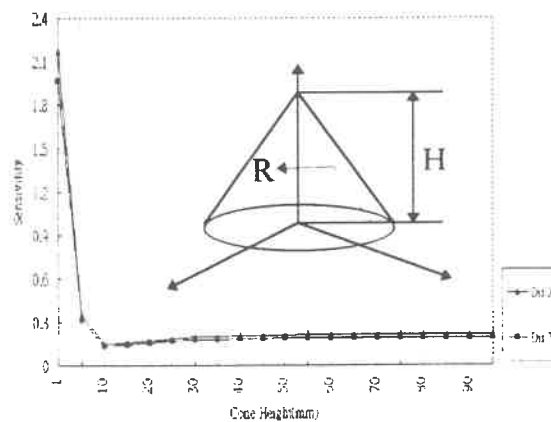


Figure 4. Illustration of geometric factor on the sensitivity curves.

Uncertainty of the Coordinate Measurement Process Due to Unwanted Asperities

James G. Salsbury *

The University of North Carolina at Charlotte; Charlotte, NC 28223

Introduction

In a variety of industries, coordinate measuring machines (CMMs) are becoming more commonly integrated into the manufacturing environment. Manufacturing process control requirements are driving CMMs towards increases in measurement flexibility, speed, automation, and ease of operation. In addition, there is an increasing desire to locate the measurement process as close as possible to the actual machining process. This often results in the placement of CMMs in relatively poor operating environments. Under these conditions, the quality of the measured part surface can be negatively effected by influences such as burrs, chips, coolants, cutting oils, and dirt. These factors can create unwanted asperities [1], or "spikes" or "flyers", in the measured data which can distort the dimensional results. The impact of unwanted asperities has most often been considered controllable or negligible by proper part handling and cleaning [2]; however, as the need for faster throughput and higher accuracy measurements continue to rise, the effect of the unwanted asperities are becoming more of a concern relative to the uncertainty of the CMM measurement results.

In this paper, an approach is presented to estimate the measurement error arising from the influence of an unwanted asperity. The analysis includes the development of the relationship between the sampling technique used to measure the part feature, including the number of discrete data points, and the magnitude of the surface irregularity. This approach is then incorporated into the CMM measurement uncertainty budget in order to determine acceptable limits for the magnitude of an unwanted asperity for a given measurement process. The approach presented in this paper is developed around the measurement of a circle diameter, though the method could be expanded to other geometric features using the same techniques.

Coordinate measuring machines, like a majority of today's modern dimensional metrology instruments, measure geometric features by sampling discrete points across surfaces and then utilize mathematical fitting algorithms to determine the "substitute geometry" of the feature. The measurement method of typical CMMs, using discrete sampling and the associated fitting algorithms, has prompted research efforts and a variety of publications over the past five to ten years (for example, see [3], [4], [5]). Those issues and discussions, however, are beyond the scope of this paper; instead, the measurement methods most widely used today on CMMs were considered. These more typical methods are based on least squares fitting algorithms, smaller point data sets, and no data elimination or filtering methods. However, since continuous scanning technology is becoming more readily available today, thus making larger data sets more feasible, this study also considers larger data sets.

Measurement Error in Circle Diameters

Figure 1 shows the typical results that may be seen when an unwanted asperity influences the measurement of a circular bore. In this particular example, the large inward deviation shown in (a) was determined to be due to the measuring probe stylus contacting a small spot of coolant that had dried on the bore surface. In (b), the bore has been recleaned and remeasured. Typically, when results such as these are discovered during part measurement, a decision to either accept the data or to remeasure the part must be made. What is missing is an understanding of how the unwanted asperity affects the measurement so that limits for the magnitude of the unwanted asperity can be established.

One alternative method of handling the unwanted asperity is to eliminate it. A number of CMM manufacturers offer some type of asperity or "data flyer" elimination in their software. Some typical approaches include the use of a set threshold value [6] or by basing the threshold relative to some

* Previously with Cummins Engine Company; Fuel Systems Business; Columbus, Indiana 47201

multiple of the standard deviation of the measured points for the particular feature [7]. In both of these cases, the magnitude of the elimination value is selected by the CMM user. The elimination method may need to vary depending on the type of geometric feature being inspected, the type of machining processes used on the part, and the CMM sampling strategy chosen. Unless an intelligent approach to removing the unwanted asperity is incorporated, the software may remove data points that are actually important pieces of information regarding the surface of the feature.



(a) Diameter = 19.990 mm, Roundness $\cong 0.070$ mm

(b) Diameter = 20.001 mm, Roundness $\cong 0.010$ mm

Figure 1: Influence of an unwanted asperity on the measurement of an internal diameter (measured with 12 points).

To develop the relationship between the magnitude of the unwanted asperity and the error in the measurement of the circle diameter, a symmetrical set of n data points around a complete circle is considered. Each point ($i = 1$ to n) is represented by the coordinate point (r_i, θ_i) . Using a least-squares algorithm, the diameter, D_o , can be expressed as follows [8]:

$$D_o = 2 (\sum r_i) / n \quad (1)$$

If we now introduce an unwanted asperity at one of the data points ($i = 1$) with a magnitude, δ , that is significantly higher than the form error already present in the original data points, the measured diameter of the circle, D_m , can be approximated as follows:

$$D_m \cong 2 [(\sum_{i=2}^n r_i) + (r_1 + \delta)] / n \quad (2)$$

$$D_m \cong D_o + (2 \delta / n) \quad (3)$$

Therefore, the change in the diameter, ΔD , due to the magnitude of the unwanted asperity, δ , is defined as follows:

$$\Delta D \cong 2 \delta / n \quad (4)$$

Equation 4 can be expanded to include the influence due to multiple unwanted asperities, δ_j , as follows:

$$\Delta D \cong 2 (\sum \delta_j) / n \quad (5)$$

It must be noted that in practical applications, as will be discussed later, the magnitude of δ can be difficult or even impossible to determine. Equations (2) to (5) are therefore shown as approximations.

Experimental Results from a CMM

The validity of equation (4) was tested in simulation using a typical CMM's software. In the simulation, a circle with a known and set diameter D_o was "probed" with n theoretically perfect points. One of these probing points was then set to have coordinates that would simulate an unwanted asperity with magnitude δ . In the simulation, the outside diameter, D_o , was set at 20 mm and δ was set at 0.100 mm. The resulting change in the diameter as a function of the number of points is shown in Figure 2, and as expected, the results follow equation (4). Using this approach, therefore, it is possible to estimate the measurement error due to the introduction of an unwanted asperity with a known magnitude. The goal of this study, however, is to be able to look at a completed measurement and estimate the measurement error due to a visible unwanted asperity. In practical situations, one has to be able to determine the magnitude of the unwanted asperity before the equations established above can be of any value.

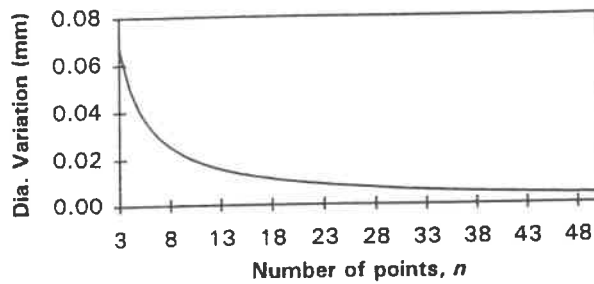


Figure 2: Change in the diameter, ΔD , as a function of the number of measurement points n .

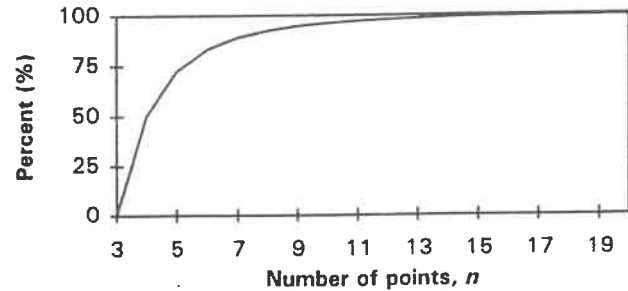


Figure 3: Percent of the unwanted asperity magnitude seen by the measurement as a function of the number of measurement points n .

In the process of doing the simulations, the measured form of the circles was also recorded. As shown in Figure 3, as the number of data points decreases, less of the magnitude of δ is seen by the measurement process. In other words, when fewer data points are taken, it becomes increasingly more difficult to know the "true" magnitude of the unwanted asperity. To measure more than 90% of the magnitude of δ , at least eight measurement points must be used. With sixteen or more points, over 99% of the magnitude is measured. If under eight points are used, the percentage drops very rapidly. In the extreme case, if only three points are used, the measured form obviously drops to zero (as three points define a circle). As the number of measurement points decreases, it also becomes more difficult to graphically see the entire magnitude of the unwanted asperity, thereby making it difficult to determine if the form is actually due to the unwanted asperity or due to some other factor, such as the machining process used to make the part. It should also be noted that the probability of encountering an unwanted asperity in the data set goes down as the number of points decreases, however that probability analysis and its significance is beyond the scope of this paper.

Measurement Uncertainty Budget

The previous results can be incorporated into the uncertainty budget for the CMM measurement process. By starting with a known and documented measurement process with an established uncertainty budget, it is possible to evaluate how much additional uncertainty due to the influence of any unwanted asperities is acceptable. With this information, a limit for the magnitude of any unwanted asperities can be determined. The following approach to estimating measurement uncertainty is in accordance with the *Guide to the Expression of Uncertainty in Measurement (GUM)* [9]. Table 1 shows an uncertainty budget for the measurement of a bore diameter with a high accuracy CMM that was developed in previous work [10]. As shown, the expanded uncertainty for the measurement process is 1.32 μm .

Table 1: Uncertainty Budget: Bore Diameter.

Source of uncertainty	Standard uncertainty
Machine	0.26 (Type B)
Stylus size	0.58 (Type B)
Temperature difference between part and machine	0.08 (Type B)
Measurement process repeatability	0.17 (Type A)
Combined standard uncertainty: $u_c = 0.66$ micrometers (μm)	
Expanded uncertainty (with coverage factor, $k = 2$): $U = 1.32$ micrometers (μm)	

As an example, assume that the acceptable uncertainty of this CMM measurement process has been established as $\pm 2.00 \mu\text{m}$ based on the quality plan of the manufacturing process. By dividing this by the coverage factor of two, the acceptable combined standard uncertainty is determined to be $1.00 \mu\text{m}$. From this, the acceptable standard uncertainty for the influence of an unwanted asperity, u_a , can be calculated as follows:

$$u_c = [(0.26)^2 + (0.58)^2 + (0.08)^2 + (0.17)^2 + u_a^2]^{1/2} \mu\text{m} \quad (6)$$

$$u_a = 0.75 \mu\text{m} \quad (7)$$

The standard uncertainty value can then be converted into a limiting value for the maximum change in the diameter, ΔD . It is assumed that the magnitude of the effect of the unwanted asperity is equally likely to be any value up to $\pm \Delta D$. This implies a rectangular or uniform distribution for the values. Using this model, the limiting value of ΔD can then be approximated from the standard uncertainty value as follows [9]:

$$u_a \cong \Delta D / (3)^{1/2} \quad (8)$$

$$\Delta D \cong (0.75) * (3)^{1/2} \cong 1.30 \mu\text{m} \quad (9)$$

This maximum value of ΔD is then used in equation (4), which results in the following expression for the maximum allowable magnitude of an unwanted asperity, δ , for a given number of data points, n :

$$\delta = 0.65 n \mu\text{m} \quad (10)$$

Using equation (10), control limits for the magnitude of δ can be established to ensure that the uncertainty of the measurement process does not exceed acceptable limits due to larger unwanted asperities. Additionally, this equation can be used to help decide upon an optimal sampling density. For example, if ten inspection points are used for this circle, the control limit for δ is only 6.5 μm ; however, if 50 points are taken, the control limit for δ increases to a fairly large 32.5 μm . The impact of this information can then be weighed against the time required to sample the additional points, the time and cost of measurement and potential remeasurements, the probability of probing unwanted asperities, and other factors influencing the sampling strategy for the particular feature.

Conclusions

Measurement can be a slow and costly operation. New CMM technology is allowing for higher data density sampling, particularly through the use of high-speed continuous scanning, which allows more of the part surface to be measured. With this trend, the occurrence of unwanted asperities in data sets may become more prevalent. Unless more intelligent methods are developed to sort out and eliminate unwanted asperities from the actual part surface data, CMM users will have to find methods to deal with the potential errors caused by the unwanted asperities. The cost of recleaning and remeasuring parts is often very high. This paper has shown that it is possible to quantitatively determine the effect of the unwanted asperities on the measurement results. The understanding of this relationship can then be used when developing an uncertainty budget for the CMM measurement process. The result is the definition of control limits for the unwanted asperities that will indicate to the CMM user when the data are questionable and when remeasurement of the part is recommended.

References

1. Term "unwanted asperity" was suggested by Mark Malburg, Cummins Engine Company, 1996.
2. Nilsson, J. T., "Application Considerations," *Coordinate Measuring Machines and Systems*, Bosch, J. A. (ed.), Marcel Dekker, New York, NY, 1995.
3. Hocken, R. J., Raja, J., and Babu, U., "Sampling Issues in Coordinate Metrology," *Manufacturing Review*, Vol. 6, No. 4, December 1993.
4. Hopp, T., "Computational Metrology," *Manufacturing Review*, Vol. 6, No. 4, December 1993.
5. Phillips, S. D., "Performance Evaluations," *Coordinate Measuring Machines and Systems*, Bosch, J. A. (ed.), Marcel Dekker, New York, NY, 1995.
6. CONCERTO User's Reference Manual, SIP, Geneva, Switzerland.
7. SCEPTRE Software Manual, Electronic Measuring Devices, Flanders, New Jersey.
8. Lin, S., Varghese, P., Zhang, C., and Wang, H. P. B., "A Comparative Analysis of CMM Form Fitting Algorithms," *Manufacturing Review*, Vol. 8, No. 1, March 1995.
9. International Organization for Standardization, *Guide to the Expression of Uncertainty in Measurement*, Geneva, Switzerland, 1995.
10. Salsbury, J. G., "A Simplified Methodology for the Uncertainty Analysis of CMM Measurements," Society of Manufacturing Engineers Technical Paper IQ95-155, 1995.

THREE-DIMENSIONAL METROLOGY FRAME FOR PRECISION APPLICATIONS

Maurice J. Prather, James F. Cuttino, and Dale E. Schinstock
Department of Mechanical Engineering
The University of Alabama

Abstract

A unique, three-dimensional, non-contacting metrology frame is presented. The metrology system uses information from three orthogonal XY stages to produce a six degree of freedom spatial representation of the object being measured. The forward and inverse kinematic solutions are presented and subsequently used to produce both a working envelope and an error budget for the metrology system.

Introduction

Complex, three-dimensional mechanical devices such as hexapod milling machines are faced with the challenge of measuring motion in six degrees of freedom. In the case of hexapods, encoders are often used on the motors of the actuators or linear sensors are mounted on the legs to measure the motion of the platform. This form of measurement directly couples the metrology system to the object being measured and subjects it to the same errors to which the device is subjected.

A three-dimensional, non-contacting metrology system (3DMS) is presented which provides position and orientation information while physically decoupling the metrology frame from the object being measured. By decoupling the metrology system from the mechanism, measurement errors are reduced and the design of the mechanism can be greatly simplified.

Conceptual Design

As shown in Figure 1, the 3DMS utilizes three orthogonal stages to produce the 6 degree of freedom (d.o.f.) representation of the object being measured. Three orthogonal laser

diodes are mounted to the object and project their beams onto the orthogonal stages of the metrology system. Mounted on each XY stage is a four quadrant photo detector that measures the variance of the laser beam from the detector's center point. As the body being measured moves through space, the corresponding XY stages track the movement of the laser beams using the photo detector signals as feedback. In turn, the position of each XY stage is measured using plane mirror laser interferometers. This information is then used to provide a solution for the 3-D spatial representation of the object.

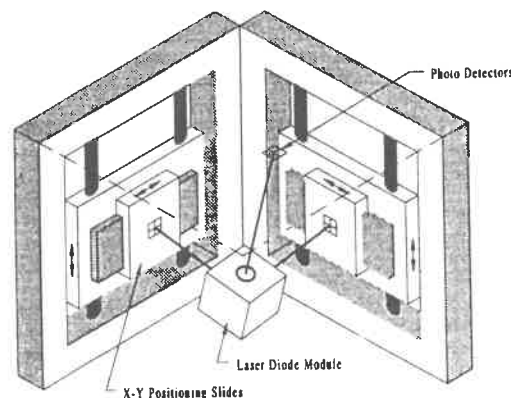


Figure 1: Metrology Frame (Top stage and interferometers removed for clarity)

Kinematic Solution

Both the forward and inverse kinematic solutions have been developed for the 3DMS. In the inverse kinematics it is assumed that the position and orientation of the body are known in the global coordinate system. Using this and known information about the construction of the system, the XY coordinates of each detector can be determined in its respective stage frame. In contrast, the forward kinematic solution assumes that the detector

coordinates are known in their respective stage frames. This information is used to determine spatial representation of the body in the global metrology frame.

The development of both kinematic solutions uses five coordinate frames. One frame is associated with each XY stage, one is attached to the body being measured, and one is a global coordinate system to which all other frames are referenced. The paths of the laser beams are nominally coincident with the axes of the body's reference frame and are referred to as the diode vectors.

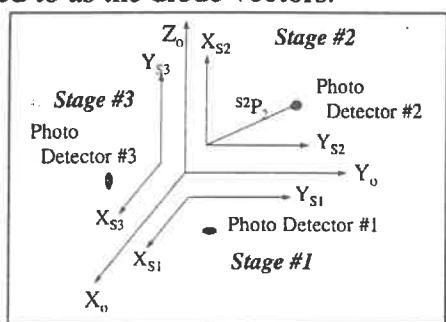


Figure 2: Axes definition

Inverse Kinematics

In the inverse kinematics the known position and orientation of the body (b) relative to the global origin (o) are represented by the homogeneous transformation matrix, oT_b . The location of each detector with respect to its corresponding stage frame (s_1 , s_2 , or s_3) is given by,

$${}^{si}P_i = {}^{si}T_o {}^oT_b {}^bP_i \quad (1)$$

where i represents both the stage and detector number as shown in Figure 2. The transformation matrices, ${}^{si}T_o$, are constructed using physical measurements of the metrology system. The position vector from the body to a corresponding photo detector, bP_i , is unknown, but it can be separated into a known unit vector and an unknown magnitude,

$${}^bP_i = m_i {}^b\hat{p}_i \quad (2)$$

Substituting (2) into (1) results in three scalar equations that are linear in the magnitude and are represented by the vector equation

$${}^{si}P_i = ({}^{si}R_o {}^oR_b {}^b\hat{p}_i) m_i + {}^{si}R_o ({}^oP_b - {}^oP_{si}) \quad (3)$$

Here, ${}^{si}R_o$ and oR_b are the rotation matrices from the stage frame to the global coordinate system and from the global coordinate system to the body, respectively.

The scalar equation in Equation 3 corresponding to the z component can be used to solve for m_i , since the z component of the detector is known in its stage frame. Substituting m_i back into the other two equations gives the XY coordinates of the detector, completing the inverse kinematics.

Forward Kinematics

The forward kinematic solution begins by equating the measured detector locations, oP_i , to the unknown diode vectors through a transformation.

$${}^oP_i = {}^oT_b {}^bP_i \quad (4)$$

This transforms the diode vector from the body reference frame to the global coordinate system. Subtracting any two of the three position vectors yields

$${}^oP_i - {}^oP_j = {}^oR_b [m_i {}^b\hat{p}_i - m_j {}^b\hat{p}_j] \quad (5)$$

Equation (5) contains two unknown scalar values, m_i and m_j , and an unknown rotation matrix oR_b . Using (5) and the fact that rotating a vector does not change its magnitude, the following equation can be written:

$$|{}^oP_i - {}^oP_j|^2 = |m_i {}^b\hat{p}_i - m_j {}^b\hat{p}_j|^2 \quad (6)$$

Equation (6) represents a triangle drawn between the body and two of the detectors. Using the law of cosines and assuming that the two unit vectors are orthogonal, then

$$\left| m_i {}^b \hat{p}_i - m_j {}^b \hat{p}_j \right|^2 = m_i^2 + m_j^2 - 2 m_i m_j \cos(\theta_{ij}) \quad (7)$$

$$\left| {}^o P_i - {}^o P_j \right|^2 = m_i^2 + m_j^2 \quad (8)$$

Three combinations for i and j are used in (7) and (8) to obtain three equations. These equations are solved for the three magnitudes. The rotation matrix is then calculated by solving equation (5). Using the rotation matrix, the Euler angles describing the orientation are found. The position vector of the body is calculated using the equation

$${}^o P_b = {}^o P_i - {}^o R_b m_i {}^b \hat{p}_i. \quad (9)$$

This completes the forward kinematics solution with the assumption that the three laser diodes are orthogonal.

Error Budget

An error budget was developed for the 3DMS using the following procedure. First, the spatial position and orientation of the body are assumed. The error components (such as non-orthogonalities) are then introduced into the system, and the inverse kinematics are used to find the "actual" locations of the detectors. Since the metrology system will at this point be unaware of error motions, the predicted position and orientation of the body are calculated using the forward kinematics and assuming no kinematic errors. The final error budget is determined by the differences between the assumed and the predicted spatial representations.

Two errors are incorporated into the inverse kinematics, non-orthogonality of the stages and non-orthogonality of the laser beams. Two errors are incorporated after the inverse kinematics by modifying the detector coordinates obtained. These two errors are detector motion in the z direction of each stage and effects due to yaw. The z motion error of each stage is comprised of three components, pure z motion of the stage and error due to roll and pitch motions. Other error motions in the

stages do not affect the error budget since they are measured by the interferometers.

The specific 3DMS being studied consists of air bearing stages with 4" travel. The vertical and angular error values were obtained from the stage manufacturer and are $0.5 \mu\text{m}$ and 3 arc-seconds (roll, pitch, and yaw), respectively. The non-orthogonalities of the laser beams and the XY stages were assumed to be 5 arc-seconds but will depend on the physical construction of the 3DMS.

Determining the error over the entire workspace of the 3DMS provides useful insight to the interaction between the stages and the body. The error budget in Figure 3 shows the position error when the object is translated with no rotation. The error was mapped across the horizontal workspace of the 3DMS at four equally spaced heights. The upper left graph is the minimum z position. In sequence from left to right, the height is increased to the maximum z position. The error magnitude increases as the body moves away from each of the stages.

Orientation error for the pure translation case is shown in Figure 4. The maximum orientation error is 8 arc-seconds and occurs when z is maximum. The non-symmetry is due to the fact that stage 2 is tilted towards stage 1 while stage 3 is tilted away from stage 1 in the x and y directions, respectively.

Figures 5 and 6 show the position errors and orientation errors for translations with 10° rotations in the x , y , and z axes. The position error for both cases changes less than 30% with the additional rotations. As with the position error budgets, the orientation error magnitude in Figures 4 and 6 increases as the body's height is increased; however, the error magnitude plateaus toward the extreme corner of the 3DMS. This plateau is indicative of an increased angular resolution due to increased radial length of each laser beam.

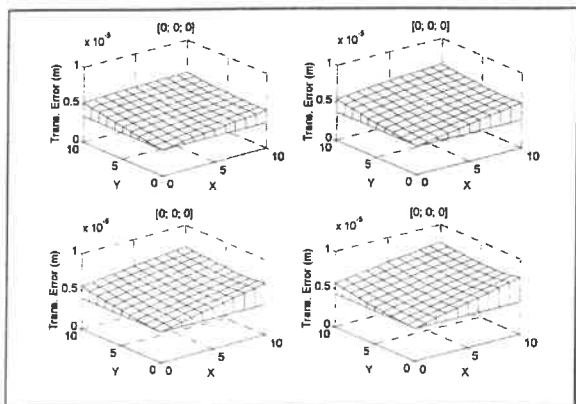


Figure 3: Position Errors for Pure Translation

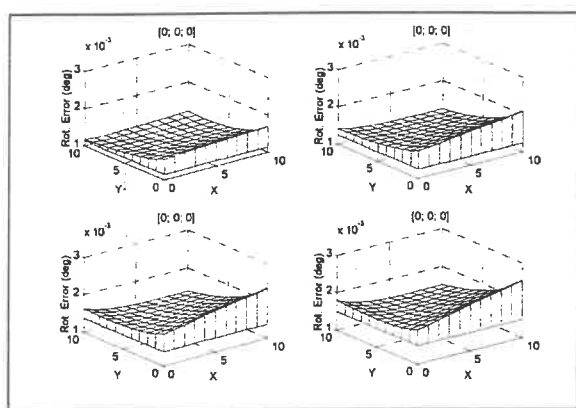


Figure 4: Orientation Errors for Pure Translation of the Object (no rotation)

Workspace Envelope

The inverse solution was used to calculate the working envelope of the 3DMS. A 2-dimensional grid representing a single XY stage was used to determine the maximum yaw angles that can be measured. Figure 7 shows the relationship between the maximum yaw angle and the working envelope for the single stage.

Results and Conclusions

A unique, non-contacting metrology system, the 3DMS is introduced. The forward and inverse kinematic solutions for the 3DMS are presented. The process by which error motion values are incorporated into the overall solution is also described. The error budget for the 3DMS indicates position accuracy within 10 micrometers is easily obtainable.

Rotational accuracy within 10 arc-seconds or less is possible depending upon the location of the body within the workspace.

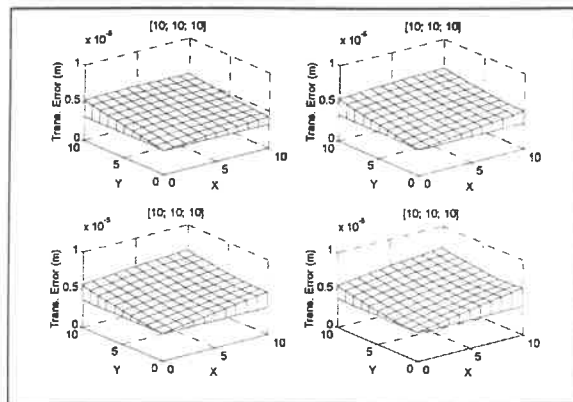


Figure 5: Position error for Translation of the Object and Rotations of θ , Φ , and $\Psi = 10^\circ$

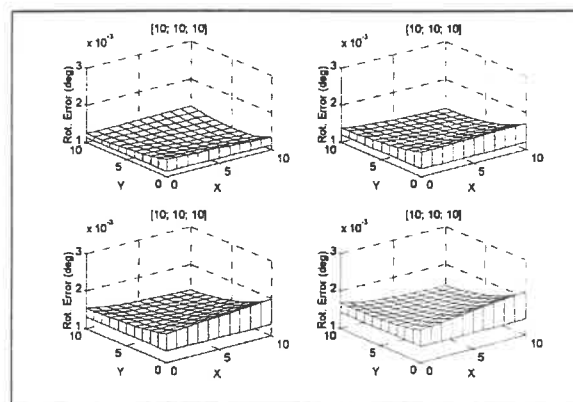


Figure 6: Orientation error for Translation of the Object and Rotations of θ , Φ , and $\Psi = 10^\circ$

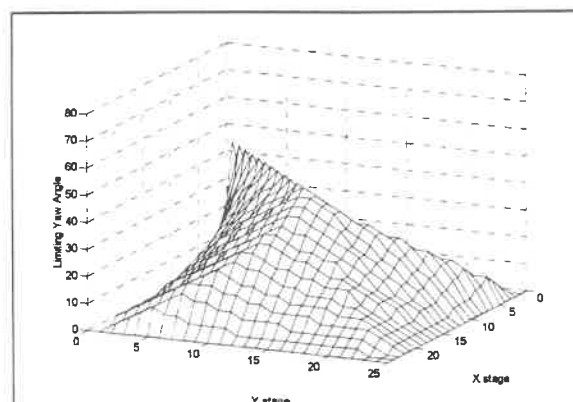


Figure 7: Limiting yaw angle

ASSESSMENT OF PLATEAU HONED SURFACE TEXTURE USING WAVELET TRANSFORM

Xiangyang Liu, Jay Raja and Harikrishna Sannareddy
Precision Engineering Laboratory
Department of Mechanical Engineering and Engineering Science
The University of North Carolina at Charlotte
Charlotte, NC 28223, USA

Introduction

Plateau honing is an important finishing technique used in the production of cylinder bores. Plateau honed surface is an "engineered surface" using two processes, namely rough and fine honing, in one setup. Monitoring and controlling the plateau honing process can be accomplished by characterizing these multiscale features comprehensively. The assessment and characterization of plateau honing and other multiprocess surface is difficult using traditional surface texture analysis. This paper briefly reviews the techniques used for characterizing honed surfaces and presents the application of wavelet transform to characterize plateau honed surface.

Analysis of plateau honed surfaces

Plateau honing is used to generate a particular type of surface structure that consists of deep grooves for oil retention and fine plateaus for high bearing capability. The plateau honed surface is produced by first rough honing to generate deep scratches and the then fine honing to remove the peaks and obtain a plateau of fine machine marks. The resulting surface texture is a combination of the marks with various scales of amplitudes and wavelengths generated by the rough and fine honing process. A typical plateau honed profile is shown in the figure 1.

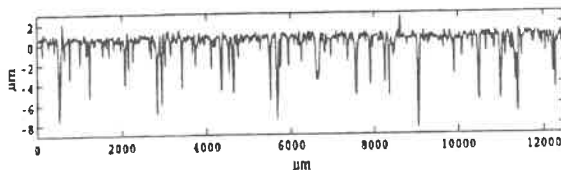


Figure 1 Plateau honed surface texture

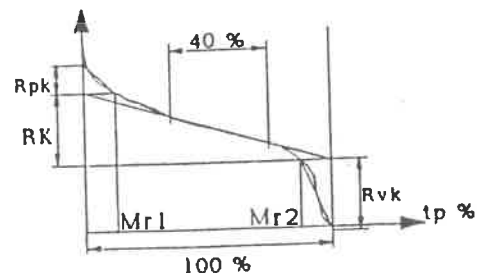


Figure 2 R_k parameters

A common method for characterizing plateau honed surface is using R_k parameters, which is specified according to DIN 4776 [1]. In this approach, the original plateau honed surface texture is divided three regions namely, extreme peaks, core and extreme valleys, shown as in figure 2. First, a minimum slope line is calculated

obtained within a 40% t_p window and the core of the profile is designated. Subsequent computation of parameters based on this line and the regions above and below it result in the following parameters such as kernel or core roughness depth R_k , reduced peak height R_{pk} , reduced valley depth R_{vk} , material ratio M_{r1} and M_{r2} . M_{r1} and M_{r2} establish a measure of the bearing ratio at the depth where the profile core meets the peaks or valleys respectively.

Another method for characterizing plateau honed surfaces is based on analyzing the cumulative distribution plot on a normal probability graph^[2, 3]. A Gaussian distribution of each component of the plateau honed surface texture is appropriate since the abrasive grains within the honing stone are random in nature. Plateau honing can be modeled as two Gaussian processes. The plateau honing operation can be summarized by presenting the profiles, amplitude distributions and probability plots of the cumulative distribution as shown in figure 3. Figure 3a and 3b represent the individual stage of the plateau honing operation and figure 3c displays the truncation of the rough honed surface by the plateau surface. The slopes of these two lines represent the R_a values of the fine and rough texture. This information can be used to control the fine and rough honing processes.

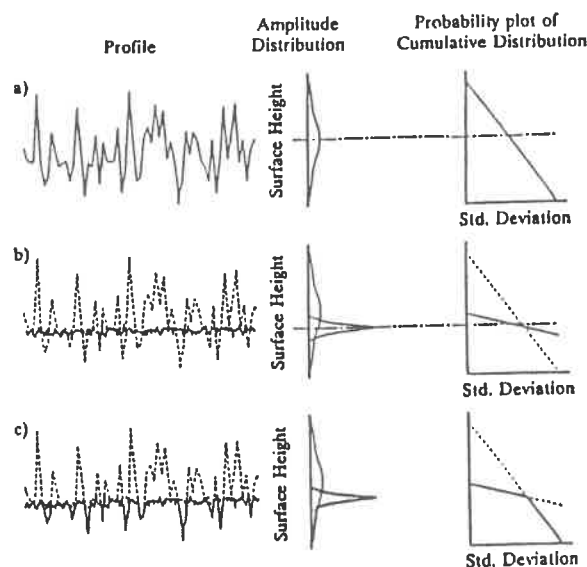


Figure 3 Characterization based on probability model

Wavelet transform for characterizing plateau honed surfaces

The wavelet analysis is performed using the basis family functions derived from the mother wavelet function by translation and dilation. The wavelet basis constructed this way has a flexible time-frequency window localized on time-frequency plane^[4]. This provides a set of windows that change in proportion to the scale of the local event

observed and whose frequency bandwidth matches the frequency of the component studied. The fundamental idea behind wavelet transform is to analyze signal according to scale, i.e. process data at different scale or resolution, shown as in figure 4. Relatively large structures of a signal is looked at a low resolution and then smaller structures are added into the interpretation by increasing the resolution. This method appears to have the potential for analyzing the multiscale aspects involved in plateau honed surfaces. The property of good time-frequency localization with a flexible time-frequency resolution of wavelet transform is explored for characterizing honed surface. The Daubechies wavelet basis is used in this study.

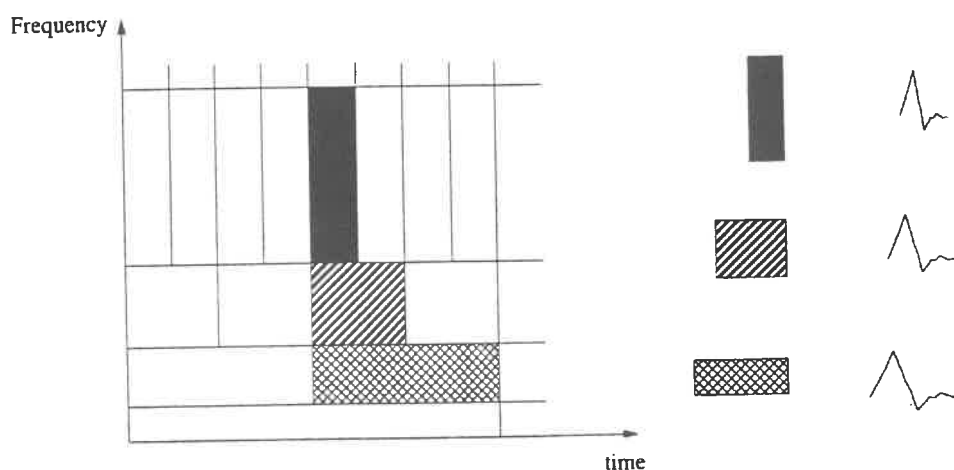


Figure 4 Flexible time-frequency resolutions of wavelet transform

In order to characterize the plateau honed surface texture, the wavelet coefficients are computed first and then the wavelet coefficients are divided into two groups based on the ratio of the valley energy to the total energy of the profile. The selection of this ratio will determine how well we separate the valley region and the plateau region. Generally, the key features of the texture are involved in the wavelet coefficients with larger magnitude and the reconstruction of the original texture is possible by using a few wavelet coefficients of larger magnitude. In plateau honed surface texture, these coefficients contain the details of the valley region since this region is the dominant than the plateau region. For constructing valley region, all coefficients larger than the threshold are kept with full accuracy while all smaller coefficients are set to zero. For constructing plateau region, all coefficients smaller than the threshold are kept with full accuracy while all greater coefficients are set to zero. By using inverse wavelet transform, the plateau region and the valley region of the plateau honed profile can be obtained separately. The two profiles obtained after separation can be characterized separately using the traditional surface texture parameters.

Figure 5 shows the separated plateau and valley portions of a plateau-honed profile using energy ratio $R = 0.9$. The separated plateau and the valley portions shown in this figure represent the corresponding portions of the original profile. Future work is in

progress to arrive at a quantitative basis for separating the rough and fine honed texture. The wavelet transform permits the analysis of surface texture in smaller bandwidths. These smaller bandwidths could be related to process parameters.

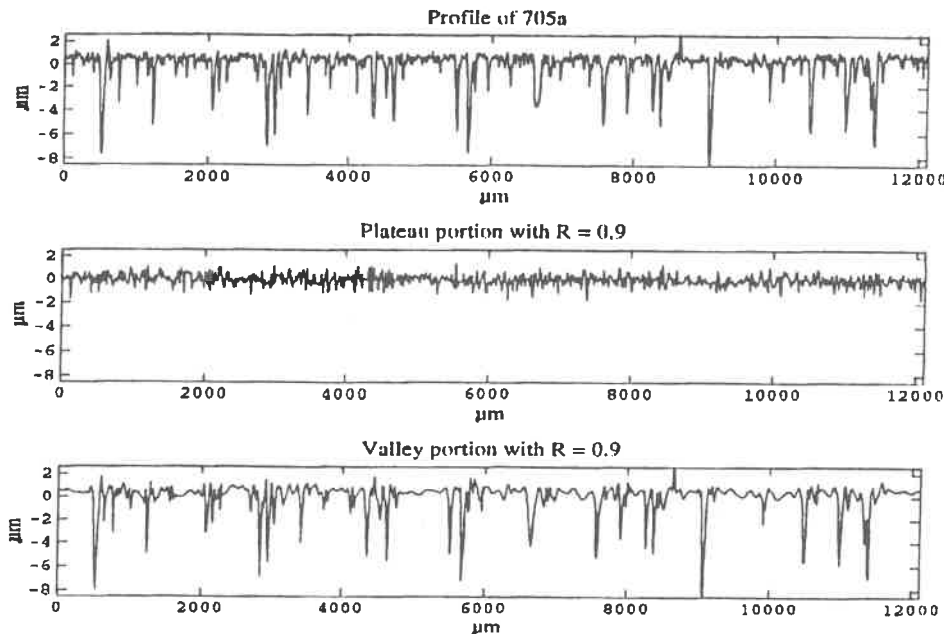


Figure.5 Separation of plateau honed surface texture by wavelet transform

References

1. DIN 4776, 1990, Measurement of surface roughness.
2. Mark, C. Malburg, Jay Raja, "Characterization of Surface Texture Generated by Plateau Honing Process", Annals of the CIRP Vol. 42/1/1993.
3. Mark. C. Malburg, 1989, The characterization of Surface Texture Generated by Multi-Process Manufacture, M. S. Thesis.
4. M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design", IEEE Transactions on Signal Processing, Vol. 40, 1992.

Authors Index

Abackerli, Alvaro J.	223
Abadin, Zainul	638
Ahn, Hyeong-Joon	328
Ahn, Jung-Hwan	496
Aketagawa, Masato	150
Allen, James P.	648
Anneccchino, L.	566
Arcona, Christopher	412, 500
Ascanio, Gabriel	346
Atherton, Paul D.	20
Badami, Vivek G.	391
Baines, R. W.	467
Baker, R.	80
Baldwin, Jon M.	654
Barnes, C.	432
Bhaskar Dixit, C. S.	551
Bibler, Jared E.	62
Bierden, Paul.	69
Bifano, Thomas G.	69, 532
Blaedel, Kenneth L.	62, 400
Bowen, D. Keith.	363
Brinksmeier, Ekkard	56, 506
Brown, C. W.	654
Caggiano, Helen E.	69, 532
Cai, Guang Qi	578
Campos Rubio, J. C.	294
Carlisle, Keith	51
Cassou, R. M.	654
Castle, Philip J.	154
Chang, In-Bae	328
Chang, S. B.	416
Chee, W. K.	490
Chen, Jenq-Shyong	273
Chen, Jun	132
Chen, Xuemei	520
Chen, Y.	51
Chetwynd, Derek G.	141, 336, 363
Chu, Xiaohua.	642
Chun, Keyoung Jin	300
Cook, Lee	80
Cuttino, James F.	310, 668
Daniel, Jay.	510, 561
Danyluk, Steven	80
Darnold, Joel R.	15
Davis, L. A. J.	141
Day, Robert D.	432
de Groot, Peter	350
DeBra, Daniel B.	114
Díaz, Rufino.	158
DiGiacomo, Benedito	223
Dixson, Ronald	212, 375

Dong, Weiping.	566
Dornfeld, David A.	496, 510, 520, 561
Dow, Thomas A.	412, 432, 500
Downs, Michael J.	91
Duduch, Jaime G.	294
Edgeworth, Robert	229
Enami, Toshio	86
Enke, Daniel	306
Entwistle, Rodney D.	420
Ertekin, Yalcin M.	278
Falter, Karl J.	412
Fang, Fengzhou	95
Ferreira, Placid M.	638
Finlayson, C. Jeffrey	310
Fischer, S.	51
Fon, Tien-Ron	104
Ford, D. G.	648
Frank, Thomas	36
Fukada, Shigeo	314
Furutani, Katsushi	86
Galindo, E.	346
Gao, Wei	163, 186
Garrard, Kenneth P.	432
Gatzen, Hans H.	536
Ge, Ping	396
Gee, Anthony E.	190, 294
Gleason, Jr., Eugene A.	167
Goto, Takayuki	146
Gowri, S.	583
Gutierrez, Hector M.	424
Gweon, Dae-Gab	626
Hahn, P. O.	381
Han, Dong-Chul.	328
Harb, Salam M.	41
Hasche, Klaus	369
Hashiguchi, Makoto.	446
Hashimoto, Hiroshi	602
Hatai, Tadaaki	614
Hatakeyama, Minoru	428
Hatamura, Yotaro	74, 476, 515
Hatch, Douglas J.	432
Hazra, Lakshman	172
Helmreich, Dieter.	381
Henderson, David	318
Henke, Richard P.	654
Hicks, Thomas R.	20
Higashi, Daisuke	176
Hirao, Masatoshi	472
Hocken, Robert J.	322
Hogan, Shawn A.	47
Holmes, Michael L.	322

*Extended abstract not available at press time.

*Abstract not available at press time.

Holster, P. L.	51	Liu, X-M.	578
Howard, Lowell P.	357, 385	Liu, Xiangyang.	672
Huang, Pei Sen	644	Ma, De Long	578
Ide, A.	482	Ma, Evan	122
Iida, I.	602	Ma, Ping.	+
Ikawa, Naoya.	65	Ma, Quanhe.	642
Imai, Kenichiro.	602	Maetzig, J. Chris.	536
Inamura, Toyoshiro	65	Masreliez, Karl G.	47
Inoue, Yasuyuki	442	Matsumoto, Kiyoshi.	476
Ishikawa, Ken-ichi.	486, 542, 572	Matsunaga, Mitsushi.	442
Jablonski, Richard.	100	McCarthy, Michael B.	190
Jared, Bradley H.	432	McConnell, Malachy	20
Jenkins, Hodge E.	406	McWaid, Thomas H.	375
Jensen, Dana	318	Mehrotra, Deepak.	237
Jokiel, Jr., Bernhard	233	Mertens, B. M.	26
Jordan, Jay D.	15	Mess, Francis	80
Jouaneh, Musa	396	Miller, Michele H.	589
Joung, Yoon	328	Mitobe, Kazuhisa.	546
Jusko, O.	369	Mitsui, Kimiyuki.	146
Kallenbach, Eberhard	456	Mitsuishi, Mamoru.	446, 515
Kamiya, Kazuhide	110, 353	Miyashiro, Hiroshi	110, 353
Kanai, Akira.	614	Miyashita Masakazu	614
Kaneeda, Toshiaki.	176, 381	Mizuhara, Kiyoshi	243
Kasaishi, S.	381	Mohri, Naotake	86
Kato, Hideo	172	Monterio, A. F.	336
Kato, Shigeo	332	Moon, Joon Hee	284
Kawabe, Shigeru.	243	Moon, Kee S.	263
Kerstan, Michael	381	Moon, Sungho.	247
Khan-Malek, Chantal.	122	Moorefield II, George M.	412, 432
Kikuri, N.	32	Morita, Noboru.	606
Kim, Hee Sool	436	Mostafavi, M. Taghi.	196
Kim, Oui-Serg	328	Murai, Hisao	381
Kim, Young Sam	284	Murata, Ryoji.	342
Kiran, M. B.	180	Nagao, Takaaki.	446, 515
Kiyono, Satoshi.	163, 186	Nakagawa, Tatsuo.	486
Ko, Tae Jo	436	Nakamura, Ryohei	172
Kobayashi, Keiko	150	Nakao, Masayuki.	74, 476
Koenders, L.	369	Nakayama, Yoshinori	150
Kothiyal, M. P.	217	Nan, Pang	602
Kouda, Nobuhiro	546	Nanjyo, Toshiyasu.	186
Koyabu, Kunio.	442	Narayanasamy, K.	583
Krylova, N. O.	363	Nasu, Yasuo	546
Kudo, Kazuki.	442	Natarajan, A.	122
Kurfess, Thomas R.	406	Nava, Rigoberto	158, 202
Kuriyagawa, Tsunemoto	557, 598	Nishida, Jun.	32
Kuroda, T.	546	Nomura, Takashi.	110, 353
Kuroda, Takaharu	172	Ohira, Fumikazu	442
Lau, Kam.	642	Okabe, Hideyuki.	446
Lêdo, H.	294	Okada, Hiroyuki.	110
Lee, Shuo-Jen.	104	Okafor, A. Chukwujekwu.	237, 278, 450
Levert, Joseph.	80	Okazaki, Yuichi	342
Li, Y.	644	Onchi, Yoshiaki.	65
Lienhard, J. H.	637*	Onnagawa, Hiroyoshi	110
Lim, Han-Seok.	496	Ozawa, Norimitsu	243
Liu Hongxiang	136	Padilla, Sergio.	202

Pahk, Heui Jae	284
Patiño, Juan Manuel	346
Patterson, Steven R.	391
Peirce, Scott B.	114
Pereira, Paulo H.	223
Pey, K. L.	490
Piccirilli, Peter	318
Piscotty, Mark A.	400
Pistecky, Peter V.	622
Postlethwaite, Scott R.	648
Potzick, James	206
Pragalsingh, Satish	196
Prather, Maurice J.	668
Preuss, Werner	56, 506
Qu, Weimin	589
Radhakrishnan, V.	180
Raja, Jay	672
Ramamoorthy, B.	180
Rehsteiner, F.	288
Rhorer, Richard L.	412
Riemer, O.	506
Rivera, G.	432
Ro, Paul I.	424, 462
Rosielle, P. C. J. N.	631
Ryu, Jae-Wook	626
Sadashivappa, K.	551
Saffert, E.	456
Sakuma, Itaru	172
Salant, Richard	80
Salsbury, James G.	664
Salzer, Leander J.	432
Sannareddy, Harikrishna	672
Santavy, A. P.	618
Sasaki, Shigetoshi	150
Sasaki, Shun'ichi	342
Sata, Toshio	65
Satoh, Makoto	186
Scattergood, Ronald O.	55*, 524, 593, 618
Schaefer, John	412
Schäffel, Christoph	456
Schellekens, Piet H. J.	631
Schinstock, Dale E.	668
Schmitz, Tony	118
Schmütz, J.	506
Schneir, Jason	212
Schwenke, Heinrich	167
Scire, Fred	385
Seki, Y.	482
Sekita, S.	602
Serrano-Carreón, L.	346
Sharp, Keith W.	593
Shen, Yin-Lin	247, 253
Shibata, Keiichi	546
Shimada, Soichi	65
Shin, Dae Young	300

Shirase, Keiichi	472
Shoji, Katsuo	557
Silver, R.	212
Singaperumal, M.	217, 551
Sirohi, R. S.	217
Slocum, Alexander H.	637*
Smith, Stuart T.	41, 391
Smith, Tony G.	15
Soltani, M. F.	267
Springer, Marlow E.	253
Stancil, Brent A.	462
Stejskal, Alois	357
Stenton, W. Conrad	+
Stone, Brian J.	420
Stone, Jack A.	357, 385
Sugawara, Takamitsu	163
Sugita, Naohiko	515
Sullivan, Paul J.	212, 375
Summerhays, Kim D.	654
Sutherland, John W.	263
Suwabe, Hitoshi	542, 572
Suzuki, Hirofumi	557
Suzuki, Kiyoshi	482
Suzuki, Seiki	150
Syoji, Katsuo	598
Tajbakhsh, Hooman	638
Takada, Koji	150
Takasu, N.	32
Takezawa, Nobuhiro	65
Tan, K. H.	490
Tanaka, Katsuya	446
Tang, Jianshe	520
Tanikella, B. V.	524
Tashiro, Hatsuho	110, 353
Taylor, John S.	62, 400
Teraoka, Sadakazu	486
Terashima, Shingo	606
Thiele, K.	369
Tian, GuyYun	467
Trigg, A.	490
Trumper, David L.	322
Tsai, Vincent W.	212, 375
Tsuchiya, Yoshihiro	172
Tsuda, N.	602
Uchida, N.	32
Udupa, Ganesha	217
Uematsu, Tetsuro	482
Ueno, Shigeru	428
Ukpong, Anietie	450
Uneda, Michio	542, 572
Valera, Benjamín	202
van Beek, H. Frank	26
van der Wulp, H.	26
Vázquez, R.	346
Venkatesh, Vellore C.	490

Vermeulen, J. P. M. B.	631
Vidic, Marina	41
Vieira Porto, Arthur J.	294
Vishwas, B. P.	259
Vorburger, Theodore V.	212, 375
Wang, Haibo	136
Wang, M.	+
Wang, Wanjun.	122
Wang, Xu-Jin.	598
Wang, Yiding.	263
Wasson, Kevin L.	637*
Watanabe, N.	482
Webster, John A.	566
Weck, Manfred	51, 267
Weikert, Sascha	288
Wen, Q. Y.	578
Wilhelm, Robert G.	229, 259
Wilkening, Günter	369
Williams, E. D.	212
Wronosky, John B.	15
Xu, Ying.	20
Yamada, Kazutoshi	150
Yamazaki, Kazuo.	126

Yan, Xiren	472
Yang, Ren	132
Yang, Yue	126
Yasui, Takeshi	472
Yasunaga, Nobuo	482
Yau, Hong-Tzong	273, 660
Yokomizo, Seiichi	176
Yokota, Yasuhiro	381
Yoshida, Masato	614
Yoshida, Y.	602
Yoshida, Yoshitaro.	606
Yoshikawa, Kazuo.	110, 353
You, Zheng	132
Yu, Huadong	606
Zanoni, Carl.	350
Zarudi, Irena	610
Zhang, G. X.	95
Zhang, Y.Q.	95
Zhao, X.	369
Zhao, Z. X.	467
Zheng, Huan Wen	578
Zhu, Hongxi.	136
Ziegert, John C.	118, 233